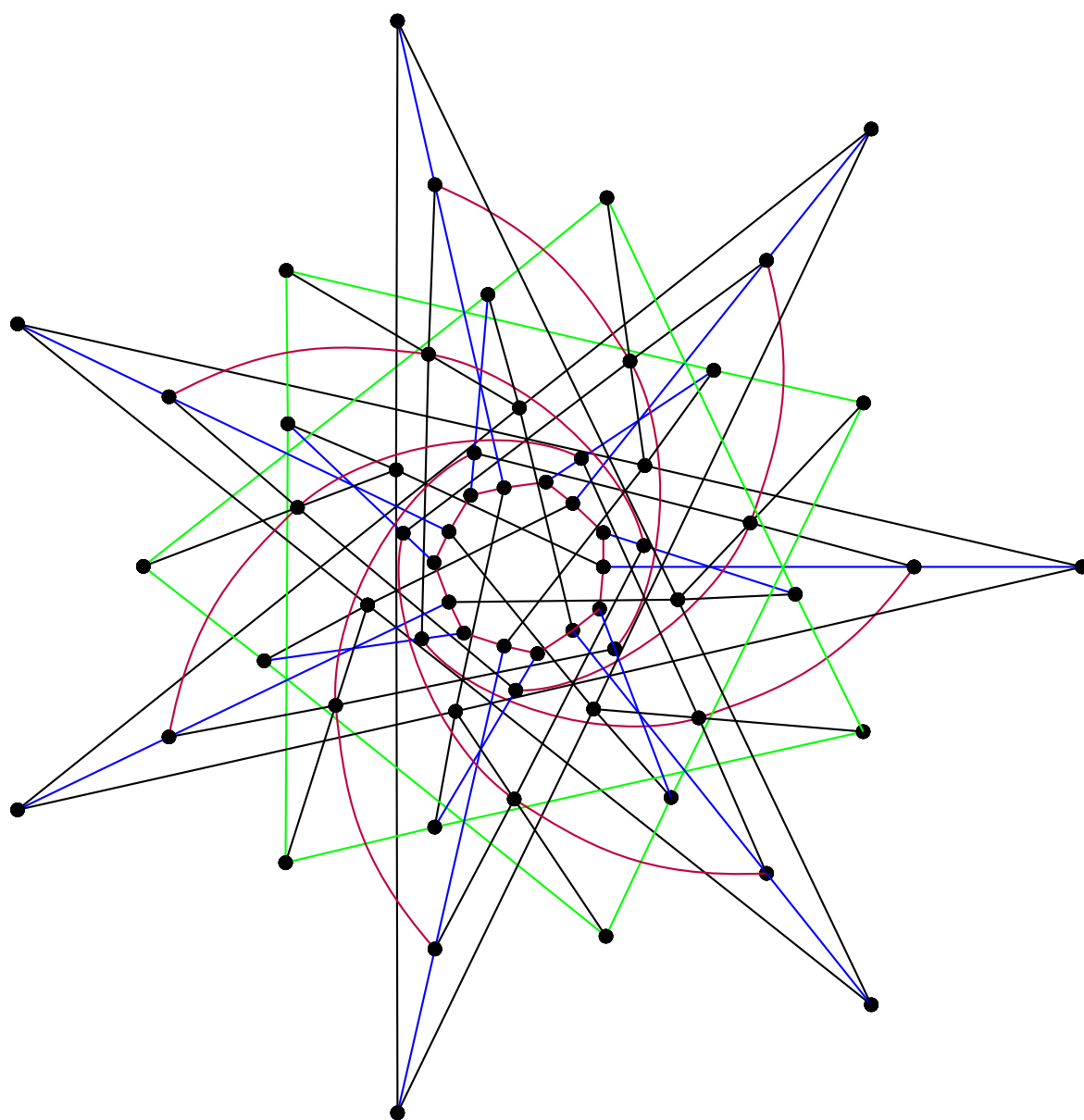


SOME CONTRIBUTIONS TO INCIDENCE GEOMETRY AND THE POLYNOMIAL METHOD

ANURAG BISHNOI



SOME CONTRIBUTIONS TO INCIDENCE GEOMETRY AND THE POLYNOMIAL METHOD

ANURAG BISHNOI

Promotor: Prof. Dr. Bart De Bruyn

A thesis presented for the degree of
Doctor of Sciences

Department of Mathematics
Ghent University
Belgium

2016-2017

Preface

The work presented in this thesis falls under two main topics, *Incidence Geometry* and *Polynomial Method*. The former deals with incidence structures that are motivated by geometrical notions like projective spaces, affine spaces and polar spaces, while the latter is a loosely defined mathematical technique of using polynomials to solve problems in combinatorics, finite geometry, discrete geometry and number theory. The thesis is accordingly divided into two parts which can be read independently of each other:

I Incidence Geometry: Chapters 1 to 6.

II Polynomial Method: Chapters 7 to 9.

The incidence structures that we will be dealing with in the first part of the thesis are *generalized polygons* and *near polygons*, while the polynomial method in the second part is related to zeros of an n variable polynomial over a ring R in certain finite subsets of R^n , which we will call *grids*.

I started my thesis work in August 2013 by learning the theory of *valuations* of generalized polygons which was developed by my supervisor Bart De Bruyn to obtain characterizations of small generalized polygons. He had successfully applied this theory to prove a spectacular result that the Ree-Tits octagon of order $(2, 4)$ is the unique generalized octagon of order $(2, 4)$ which contains a suboctagon of order $(2, 1)$. The basic idea in this theory is to first compute certain integer valued functions on the point set of a given generalized polygon, the so-called *valuations*, and a particular incidence structure formed by these valuations called the *valuation geometry* of the generalized polygon. This valuation geometry is then used to obtain the required information about every generalized polygon that contains the given generalized polygon as a full subgeometry. He saw further potential in this theory and gave me a result on semi-finite generalized hexagons, which he knew how to prove, as a “toy problem” that I could work on after mastering the techniques that he had developed. At around November 2013, I successfully solved this problem, which I consider as my first milestone in mathematical research. We had proved that (a) no (possibly infinite) generalized hexagon can contain the split Cayley hexagon $H(2)$ as a proper full subgeometry, and (b) every generalized hexagon containing the dual split Cayley hexagon $H(2)^D$ as a proper full subgeometry is finite, and hence isomorphic to the dual twisted triality hexagon $T(2, 8)$ by the classification result of Cohen and Tits. Within a couple of months we generalized the techniques further so that we could prove similar results for *near hexagons* (which are more general than generalized hexagons) containing these subhexagons. This work became my first research paper. We have recently extended our results further by considering generalized hexagons with $q + 1 \in \{4, 5\}$ points on each line containing one of the known generalized hexagons of order q as a subgeometry. All of these results constitute Chapter 4 of this thesis.

As an offshoot of our work on semi-finite generalized polygons, Bart and I thought it would be nice to compute the valuation geometry of some other well-known near polygons with

three points on each line (many of our techniques were limited to point-line geometries with three points on each line). The most natural choice for us was the Hall-Janko near octagon HJ, which corresponds to the sporadic finite simple group J_2 and contains subgeometries isomorphic $H(2)^D$. I computed the valuation geometry of HJ, and as a result we noticed that there might exist a near octagon with three points on each line containing HJ which can be constructed using certain valuations of HJ. Since there was no known example of such a near octagon, this was unexpected (and exciting!). I checked, using a computer, that indeed we have a near octagon containing HJ, and quickly started finding its properties with the help of the mathematical software SageMath. My computations revealed that the automorphism group of this new near octagon acted vertex-transitively on the collinearity graph, which was a bit surprising due to the asymmetrical nature of our construction using certain valuations of HJ. Moreover, SageMath gave me the order of the full automorphism group and the fact that its derived subgroup is a simple group. From this we conjectured that the automorphism group of our new near octagon must be isomorphic to $G_2(4):2$, where $G_2(4)$ is the well-known finite simple group of Lie type. We finally arrived at a group theoretical construction of this near octagon using central involutions of $G_2(4):2$ and proved our conjecture. Sitting inside this new near octagon, we discovered another new near octagon, which corresponds to the finite simple group $L_3(4)$ (the projective special linear group). Both of these near octagons are new and they share many common properties; for example, both of them can be defined by a suborbit diagram of central involutions which look quite similar. We give a common treatment of these two new incidence geometries in Chapter 5, where we derive several important structural properties.

The new near octagon corresponding to the group $G_2(4)$, let's call it O_1 , has HJ as a subgeometry, which in turn has $H(2)^D$ as a subgeometry. In this manner we get a “tower” of near polygons, $H(2,1) \subset H(2)^D \subset HJ \subset O_1$, where $H(2,1)$ is the unique generalized hexagon of order $(2,1)$ (it is the dual of the incidence graph of Fano plane). The automorphism groups of these near polygons are $L_3(2):2$, $U_3(3):2$, $J_2:2$ and $G_2(4):2$, respectively. Thus we have the first four members of the *Suzuki tower* of finite simple groups, $L_3(2) < U_3(3) < J_2 < G_2(4) < Suz$, where *Suz* is the sporadic simple group discovered by Michio Suzuki in 1969. Suzuki constructed *Suz* as an automorphism group of the largest graph in a sequence $\Gamma_0, \dots, \Gamma_4$ of graphs in which $\Gamma_1, \dots, \Gamma_4$ are strongly regular graphs and for each $i \in \{0, \dots, 3\}$, the graph Γ_i is the local graph of Γ_{i+1} . Using our Suzuki tower of near polygons, we can show that all of these graphs can be constructed in a uniform geometrical fashion. Moreover, we have proved that every near polygon in this tower, except the first one, is the unique near polygon of its order and diameter containing the previous near octagon as a subgeometry. These results constitute Chapter 6 of the thesis.

A significant part of my work in incidence geometry involved computations on computer models of near polygons and generalized polygons using mathematical software like GAP and SageMath. We had to design reasonably fast ways of finding certain substructures of near polygons called *hyperplanes*, which would then be used to construct the valuation geometry of a given near polygon. While working on these algorithms, I realised that that they can help us answer some open questions on distance- j ovoids of generalized polygons. In collaboration with Ferdinand Ihringer, I was able to show non-existence of distance-2 ovoids in the dual split Cayley hexagon $H(4)^D$. The techniques that we developed for this work are quite general and they might be helpful in resolving other such “small cases” of

open problems in finite geometry. All my computational work, some of which is required in subsequent chapters, is contained in Chapter 3.

Towards the end of 2014, I started getting interested in the so-called *polynomial method* after reading Terence Tao’s blog post on Zeev Dvir’s two page solution of the famous finite field Kakeya problem using a slick polynomial argument. I quickly learned about Noga Alon’s Combinatorial Nullstellensatz, Tom H. Koornwinder’s proof of the absolute bound on equiangular lines, Andries Brouwer and Lex Schrijver’s proof of the bound on affine blocking sets, and several other such important results involving (elementary) polynomial arguments. By March 2015, I started writing some notes on the polynomial method in collaboration with my friends Abhishek Khetan and Aditya Potukuchi who were also interested in understanding these techniques. We read several papers and expository articles to get acquainted with some of the basic ideas involved. In particular, we started reading the papers “Combinatorial Nullstellensatz Revisited” by Pete Clark and “Warning’s Second Theorem with Restricted Variables” by Pete Clark, Aden Forrow and John Schmitt. In April 2015, I emailed Pete and John saying that one can easily obtain a generalization of the classical Chevalley-Warning theorem (which is a corollary of Warning’s second theorem) using a result of Simeon Ball and Oriol Serra called the *Punctured Combinatorial Nullstellensatz*. My new result is in fact a generalization of David Brink’s restricted variable Chevalley-Warning theorem, which he proved using Alon’s Combinatorial Nullstellensatz. We give the exact statement in Chapter 8 where we also lay down the basics of *grid reduction* of polynomials and give a new proof of Punctured Combinatorial Nullstellensatz.

Pete and John responded promptly to my email(s) and shared with me a lot of material related to the things they have been working on. The discussions that followed with Pete, John and Aditya resulted in a collaborative work which I believe is my main contribution to the polynomial method. The main tool used by Clark, Forrow and Schmitt to obtain their generalization of Warning’s second theorem on the minimum number of common zeros of a system of polynomial equations over a finite field is a result of Alon and Füredi from 1993 that gives a lower bound on the number of *non-zeros* of a polynomial in a finite grid. They gave several combinatorial applications of their result along the general theme of refining “combinatorial existence theorems into theorems which give explicit (and sometimes sharp) lower bounds on the number of combinatorial objects asserted to exist”. Within a few weeks of our discussions, I noticed that this Alon-Füredi theorem is in fact related to Reed-Muller type affine variety codes, and to the famous Schwartz-Zippel lemma. I also came up with an alternate proof of the Alon-Füredi theorem, and then Aditya discovered an even simpler proof which helped me in obtaining a new generalization of the Alon-Füredi theorem. Ultimately, we discovered a lot more about this result and its connection with different areas of mathematics, including incidence geometry. All of this work is contained in the last chapter, Chapter 9.

Acknowledgements

I would like to thank Bart De Bruyn for being a great supervisor during the past three years. He read, and re-read, all of my work, and always gave valuable inputs. Nothing in Part I of the thesis would have been possible without him helping me identify the right problems to work on.

I would like to thank Pete Clark and John Schmitt for several helpful discussions during my PhD. My collaboration with them, which led to the second part of my thesis, has been a great learning experience.

Aditya Potukuchi deserves a special thanks for being a great friend (and a co-author) who listened to a lot my ramblings over the years. Mathematical discussions with him and Abhishek Khetan, who I would also like to thank, have been a crucial part of my PhD experience.

Another friend + co-author who deserves thanks is Ferdinand Ihringer. After my supervisor, he is probably the person who knows the most about my research work. Over the years, we discussed several problems in finite geometry, which in particular motivated me to learn some interesting topics that I was unfamiliar with. Inputs from John Bamberg in the problems that Ferdinand and I worked on were also quite helpful, and I thank him for that.

In a lot of my research work, I have used the free open-source softwares SageMath and GAP. I am grateful to the creators and developers of these mathematical softwares for their great service to the mathematical community, without whom I would not have been able to discover many of the results contained in this thesis.

I am indebted to Prof. Bhaskar Bagchi for introducing me to Incidence Geometry. My research interests have been shaped by the 2 months I spent with him as a summer research fellow at the Indian Statistical Institute, Bangalore, and I thank the Indian Academy of Sciences for giving me that opportunity. I thank Prof. Bagchi for all the valuable lessons he taught me and for informing me about the Incidence Geometry research group at Ghent University.

I can't imagine learning geometry so well without coming to this department at Ghent which has developed a great tradition in the area. Thus, I thank all the mathematicians who have contributed to the development of this research group. I would also like to thank the department secretary Samuel Perez for helping me in many different ways and handling all the paperwork.

I thank all my jury members, Bart De Bruyn, Simeon Ball, Bruce Cooperstein, Hendrik Van Maldeghem, Leo Storme and Koen Thas, for reading this thesis carefully and making valuable comments.

On a personal level, I thank my friends and colleagues who have been responsible for the wonderful time that I have had here in the past three years: Aparajita, Anwasha, Manuel, Ana, Annelies, Hendrik, Pauline, Paul (Russel), Paul (Shafer), John, Morgan, Ursula, Charlotte, Ewa, Magali, Anamari, Andrea, Frank, Geertrui, Maarten, Wouter, Marianna, Rico, Alice, Matteo, Surya, Daniele, ... Also, a special thanks to all my friends back home in India, who I visited every year and always had a great time with.

Above all, I thank my parents, my brother, and my sister for the love and support they have given me all my life.

Publications

There are seven publications related to the work presented in this thesis [18–24]. At the time of writing this thesis, three of these [19, 21, 23] have been published, three [18, 22, 24] are under review, and one [20] is under preparation.

- Chapters 3 and 4 are based on [22], [23] and [24].
- Chapter 5 is based on [20] and [21].
- Chapter 6 is based on [19].
- Chapter 8 contains an unpublished result (Theorem 8.3.1).
- Chapter 9 is based on [18].

These references and remarks will not be repeated in individual chapters.

Contents

Preface	v
Acknowledgements	ix
Publications	xi
I. Incidence Geometry	1
1. Preliminaries	3
1.1. Point-Line Geometries	3
1.2. Generalized Polygons	5
1.2.1. Important Examples	8
1.2.2. Classification of Finite Generalized Polygons	9
1.3. Near Polygons	11
1.3.1. Substructures of a Near Polygon	13
1.3.2. Near Hexagons of Order 2	15
2. Valuations of Near Polygons	17
2.1. Introduction	17
2.2. Definitions and Properties	17
2.3. Three Points on Each Line	21
2.3.1. An Example – the Doily	23
2.4. Some valuation geometries	26
3. Computing Hyperplanes, Ovoids and Valuations	31
3.1. Introduction	31
3.2. Definitions and Basic Techniques	32
3.3. Computer Models of Geometries	34
3.4. Valuation Geometries	36
3.5. Distance-2 ovoids in $H(4)^D$	43
4. Generalized Hexagons Containing a Subhexagon	47
4.1. Introduction	47
4.2. Near Hexagons of Order $(2, t)$	47
4.2.1. Near Hexagons Containing $H(2)^D$	48
4.2.2. Near Hexagons Containing $H(2)$	53
4.3. Polygonal Valuations and Intersections of Hyperplanes	55
4.3.1. Generalized Hexagons Containing Split Cayley Hexagons and their Dual	59

5. Two New Near Octagons	63
5.1. Introduction	63
5.2. Suborbit Diagrams	65
5.2.1. The $G_2(4)$ Near Octagon	65
5.2.2. The $L_3(4)$ Near Octagon	67
5.3. A Family of Near Octagons	70
5.4. Hall-Janko Suboctagons of O_1	77
5.5. The Automorphism Groups	82
5.5.1. Automorphism Group of $G_2(4)$ Near Octagon	82
5.5.2. Automorphism Group of $L_3(4)$ Near Octagon	83
5.6. Distance Regular Graphs of Soicher	85
5.6.1. The First Graph	85
5.6.2. The Second Graph	86
6. Suzuki Tower of Near Polygons	89
6.1. Introduction	89
6.2. Constructing the Graphs	91
6.3. Characterization of the Suzuki Tower	93
6.3.1. Hall-Janko Near Octagon	93
6.3.2. $G_2(4)$ Near Octagon	97
 II. Polynomial Method	 103
7. Introduction	105
8. Grid Reduction	109
8.1. Basics	109
8.2. Applications of Grid Reduction	111
8.3. Punctured Chevalley-Waring	114
9. Alon-Füredi Theorem	115
9.1. Introduction	115
9.2. Balls in Bins	117
9.3. Proof of Generalized Alon-Füredi Theorem	119
9.3.1. Sharpness of the Generalized Alon-Füredi Bound	120
9.4. Connections with the Schwartz-Zippel Lemma	121
9.5. Multiplicity Enhancements	123
9.5.1. Hasse Derivatives	123
9.5.2. Multiplicities	124
9.5.3. Multiplicity Enhanced Theorems	126
9.5.4. A Counterexample	127
9.6. Connections with Coding Theory	128
9.7. Blocking Sets and Hyperplane Coverings	130
9.7.1. Partial Coverings of Grids by Hyperplanes	130
9.7.2. Partial Covers and Blocking Sets in Finite Geometries	132
 A. Nederlandstalige samenvatting	 135
A.1. Deel 1 (Hoofdstukken 1 t.e.m. 6)	135

A.2. Deel 2 (Hoofdstukken 7 t.e.m. 9)	138
---	-----

List of Figures

1.1.	The unique GQ of order $(2, 2)$	4
1.2.	Doily with GQ $(2, 1)$ inside	5
2.1.	Lemma 2.2.1	19
4.1.	Lemma 4.2.9 (5)	51
5.1.	The suborbit diagram for the central involutions of $G_2(4):2$	66
5.2.	The suborbit diagram for the central involutions of $L_3(4):2^2$	67
5.3.	The structure of \mathcal{S} with respect to the fixed point x	75

List of Tables

2.1.	The valuations of $GQ(2, 1)$	24
2.2.	The lines of the valuation geometry of $GQ(2, 1)$	24
2.3.	The valuations of $H(2)^D$	26
2.4.	The lines of the valuation geometry of $H(2)^D$	27
2.5.	The valuations of $H(2)$	27
2.6.	The lines of the valuation geometry of $H(2)$	28
2.7.	The valuations of HJ	28
2.8.	The lines of the valuation geometry of HJ	29
5.1.	Suborbit description for $G_2(4):2$	65
5.2.	Suborbit description for $L_3(4):2$	67
5.3.	Fixed set intersections in the $G_2(4)$ near octagon	85
5.4.	Intersection patterns in Soicher's first graph	86
5.5.	Fixed set intersections in $L_3(4)$ near octagon	87
5.6.	Intersection patterns in Soicher's second graph	87

Part I.

Incidence Geometry

1. Preliminaries

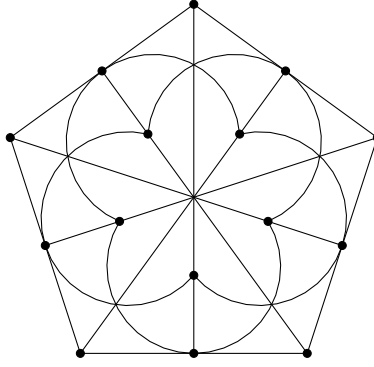
We assume a basic level of familiarity with linear algebra, projective geometry, graph theory and group theory (especially group actions and permutation groups). Here we note down some definitions and results related to the main players of the first part of this thesis, *generalized polygons* and *near polygons*.

1.1. Point-Line Geometries

A *point-line geometry* is a triple $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ of sets where \mathcal{P} is non-empty, \mathcal{P} and \mathcal{L} are disjoint and \mathbf{I} is a subset of $\mathcal{P} \times \mathcal{L}$ such that for every $L \in \mathcal{L}$ there exist at least two $x \in \mathcal{P}$ for which $(x, L) \in \mathbf{I}$. We refer to the elements of \mathcal{P} as *points* of \mathcal{S} , elements of \mathcal{L} as *lines* of \mathcal{S} and \mathbf{I} as the *incidence* relation between the points and lines. We will often use geometrical language by writing the statement $(x, L) \in \mathbf{I}$ as “the point x and the line L are incident”, “the point x lies on the line L ”, “the line L contains the point x ,” etc. Two points x, y are called *collinear* if there is a line L incident with both x and y . We say that a point-line geometry has *order* (s, t) , for possibly infinite cardinal numbers s and t , if every line is incident with $s + 1$ points and every point is incident with $t + 1$ lines. If $s = t$, then we simply say that the point-line geometry has order s . When every point is incident with at least three lines and every line is incident with at least three points, we call the point-line geometry *thick*. The *point-line dual* of a point-line geometry $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$, when \mathcal{L} is non-empty and every point is incident with at least two lines, is the point-line geometry $\mathcal{S}^D = (\mathcal{P}^D, \mathcal{L}^D, \mathbf{I}^D)$ where $\mathcal{P}^D = \mathcal{L}$, $\mathcal{L}^D = \mathcal{P}$, $\mathbf{I}^D \subseteq \mathcal{L} \times \mathcal{P}$ and $(L, x) \in \mathbf{I}^D$ if and only if $(x, L) \in \mathbf{I}$.

A point-line geometry \mathcal{S} is called a *partial linear space* if for every pair of distinct points in \mathcal{S} , there is at most one line which contains both of these points. When there is exactly one line through every pair of distinct points, we get a *linear space*. Important examples of linear spaces are *projective spaces* and *affine spaces*. When a point-line geometry $(\mathcal{P}, \mathcal{L}, \mathbf{I})$ satisfies the condition that $\{x \in \mathcal{P} \mid x \mathbf{I} L_1\} \neq \{x \in \mathcal{P} \mid x \mathbf{I} L_2\}$ for every $L_1 \neq L_2 \in \mathcal{L}$, we can uniquely identify each line with the set of points that are incidence with it, and thus we get a hypergraph (V, E) corresponding to the geometry, where $V = \mathcal{P}$ and $E = \{\{x \in \mathcal{P} \mid x \mathbf{I} L\} \mid L \in \mathcal{L}\}$. For partial linear spaces this condition is automatically satisfied; hence we can, and often will, treat a partial linear space like a hypergraph with the incidence relation being set containment.

Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a point-line geometry. The *incidence graph* of \mathcal{S} is the graph whose vertex set is $\mathcal{P} \cup \mathcal{L}$ and two vertices are adjacent when they are incident in \mathcal{S} . The *point graph* or *collinearity graph* of \mathcal{S} is the graph whose vertex set is \mathcal{P} and two vertices are adjacent when they are collinear. The distance function on pairs of vertices of the incidence graph will be denoted by $\delta_{\mathcal{S}}(\cdot, \cdot)$, while the distance function on pairs of vertices of the collinearity graph will be denoted by $d_{\mathcal{S}}(\cdot, \cdot)$. When no confusion arises we will


 Figure 1.1.: The unique GQ of order $(2, 2)$

simply write δ and d , without the subscript. We call the geometry \mathcal{S} *connected* if the diameter of its incidence graph, or equivalently the diameter of its collinearity graph, is finite. For a point x and a line L , we define $d(x, L)$ as $\min\{d(x, y) \mid y \in L\}$, and for two lines L, M , we define $d(L, M)$ as $\min\{d(x, y) \mid x \in L, y \in M\}$. For two nonempty subsets A, B of \mathcal{P} , we define $d(A, B)$ as $\min\{d(x, y) \mid x \in A, y \in B\}$. When A is a singleton $\{x\}$, we simply denote $d(\{x\}, B)$ as $d(x, B)$. Let Γ denote the collinearity graph of \mathcal{S} . Then the set of points at distance i (at most i) from a point/line/subset X in Γ is denoted by $\Gamma_i(X)$ (respectively $\Gamma_{\leq i}(X)$).¹ Unless stated otherwise, the distances measured in this thesis will be in the collinearity graph of a point-line geometry. The following proposition relates the two distance functions d and δ of \mathcal{S} . We leave its proof to the reader.

Proposition 1.1.1. *Let $(\mathcal{P}, \mathcal{L}, I)$ be a connected point-line geometry, let $\delta(\cdot, \cdot)$ denote the distance function in its incidence graph and let $d(\cdot, \cdot)$ denote the distance function in its collinearity graph. Let $x, y \in \mathcal{P}$ and $L, M \in \mathcal{L}$ with $L \neq M$. Then we have $\delta(x, y) = 2 \cdot d(x, y)$, $\delta(x, L) = 2 \cdot d(x, L) + 1$ and $\delta(L, M) = 2 \cdot d(L, M) + 2$.*

Two point-line geometries $(\mathcal{P}, \mathcal{L}, I)$ and $(\mathcal{P}', \mathcal{L}', I')$ are called *isomorphic* if there exists a bijective map $f : \mathcal{P} \cup \mathcal{L} \rightarrow \mathcal{P}' \cup \mathcal{L}'$ such that $f(\mathcal{P}) = \mathcal{P}'$, $f(\mathcal{L}) = \mathcal{L}'$ and $(x, L) \in I$ if and only if $(f(x), f(L)) \in I'$. An isomorphism from a point-line geometry to itself is called an *automorphism*, and the group formed by all automorphisms is called the *automorphism group* of the geometry.

Example. Let \mathcal{P} be the set of all two element subsets of $\{1, 2, 3, 4, 5, 6\}$ and \mathcal{L} the set of all partitions of $\{1, 2, 3, 4, 5, 6\}$ into two element subsets. Then by taking incidence relation I as set containment, we get a point-line geometry $(\mathcal{P}, \mathcal{L}, I)$ of order $(2, 2)$ with 15 points and 15 lines. This point-line geometry is isomorphic to its point-line dual and has the automorphism group isomorphic to the symmetric group S_6 . In Figure 1.1 we have given a drawing of this geometry which explains the common name given to it, *Doily* (the 15 lines are the 5 lines of the regular pentagon, 5 angle bisectors and 5 incomplete circles). As we will see in the next section, this point-line geometry is an example of a *generalized quadrangle*. In fact, this is the unique generalized quadrangle of order $(2, 2)$ up to isomorphism.

¹For any arbitrary graph Γ , we denote the set of vertices at distance i from a given vertex x by $\Gamma_i(x)$.

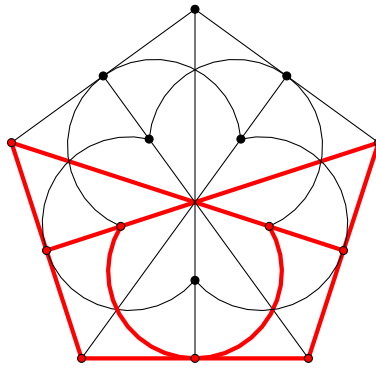


Figure 1.2.: Doily with GQ(2, 1) inside

A point-line geometry $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ is called a *subgeometry* of another point-line geometry $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathbf{I}')$ if $\mathcal{P} \subseteq \mathcal{P}'$, $\mathcal{L} \subseteq \mathcal{L}'$ and $\mathbf{I} = \mathbf{I}' \cap (\mathcal{P} \times \mathcal{L})$. The subgeometry is called *full* if for every line $L \in \mathcal{L}$, the set $\{x \in \mathcal{P} \mid x \mathbf{I} L\}$ is equal to the set $\{x \in \mathcal{P}' \mid x \mathbf{I}' L\}$. It is called *isometrically embedded* if the distances measured in the collinearity graph of \mathcal{S} are equal to the distances in the collinearity graph of \mathcal{S}' , i.e., for every pair of points x, y in \mathcal{S} , we have $d_{\mathcal{S}}(x, y) = d_{\mathcal{S}'}(x, y)$. Figure 1.2 shows a full isometrically embedded subgeometry of Doily. The subgeometry is in fact isomorphic to the unique generalized quadrangle of order $(2, 1)$. By abuse of notation, we will denote the set $\Gamma_i(\mathcal{P})$ of points of \mathcal{S}' at distance i from \mathcal{P} by $\Gamma_i(\mathcal{S})$.

The books [72], [92] and [125] are some of the standard references for point-line geometries. For an introduction to incidence geometry, see [12] and [63].

1.2. Generalized Polygons

Generalized polygons were introduced by Tits in the Appendix of his famous paper on triality [136], and they are now an integral part of incidence geometry. They have connections with several areas of mathematics like group theory, extremal graph theory, coding theory and design theory. The standard reference for generalized polygons is [138], and the older survey by Kantor [94, Section A] gives a succinct introduction to the basic theory of these point-line geometries. For proofs of statements in this section which are stated without a proof, we refer to [138] and [63, Chapter 5].

One of the easiest ways to define generalized polygons is as follows.

Definition. For an integer $n \geq 2$, a *generalized n -gon* is a point-line geometry whose incidence graph has diameter n and girth $2n$.²

A generalized polygon is a generalized n -gon for some integer $n \geq 2$. Note that a bipartite graph of diameter n which contains a cycle has girth (length of the shortest cycle) at most $2n$. Therefore, generalized polygons correspond to bipartite non-tree graphs of a given

²Sometimes the condition that every vertex of the incidence graph has degree at least 3 is also added, for example in [94]. We will simply call the generalized polygons satisfying that condition as *thick* generalized polygons.

diameter and maximum possible girth, the so-called bipartite *Moore graphs* [85, Chapter 5]. From the definition it directly follows that the point-line dual of a generalized n -gon is also a generalized n -gon. A generalized 2-gon is simply a point-line geometry whose incidence graph is a complete bipartite graph. A generalized 3-gon (triangle) is equivalent to a possibly degenerate projective plane. Generalized quadrangles as defined above can be shown to be equivalent to partial linear spaces satisfying the following axioms:

GQ1 there exist at least two disjoint lines;

GQ2 for every point x and a line L not containing x , there exists a unique point $\pi(x)$ on L which is collinear with x .

Similarly, we have the following equivalent definition of generalized n -gons for any $n \geq 3$. An ordinary n -gon for $n \in \mathbb{N} \setminus \{0, 1, 2\}$ is a point-line geometry which is isomorphic to the partial linear space that has point set $\{1, \dots, n\}$, line set $\{\{1, 2\}, \{2, 3\}, \dots, \{n-1, n\}, \{n, 1\}\}$ and set containment as incidence relation. A generalized n -gon with $n \in \mathbb{N} \setminus \{0, 1, 2\}$ is a partial linear space $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ satisfying:

GP1 \mathcal{S} has no subgeometries that are ordinary m -gons with $m \in \{3, 4, \dots, n-1\}$;

GP2 \mathcal{S} has subgeometries that are ordinary n -gons;

GP3 for every $X, Y \in \mathcal{P} \cup \mathcal{L}$ with $X \neq Y$, there exists a subgeometry of \mathcal{S} containing X and Y which is isomorphic to an ordinary n -gon.

One of the most basic properties of *thick* generalized polygons is that they have an order. We state it here without proof.

Proposition 1.2.1. *Let \mathcal{S} be a generalized n -gon for $n \in \mathbb{N} \setminus \{0, 1, 2\}$ and let $\tilde{\Gamma}$ be its incidence graph.*

- (1) *For every two vertices u and v of $\tilde{\Gamma}$ at maximal distance n from each other, there exists a bijection between the set of neighbors of u and the set of neighbors of v .*
- (2) *If \mathcal{S} is thick, then \mathcal{S} has an order (s, t) . Moreover, if n is odd, then $s = t$.*
- (3) *\mathcal{S} is thick if and only if it has ordinary $(n+1)$ -gons as subgeometries.*

For generalized quadrangles, we can say something more. An $(n_1 \times n_2)$ -grid, with $n_1, n_2 \in \mathbb{N} \setminus \{0, 1\}$, is a point-line geometry that is isomorphic to the partial linear space whose point set is equal to $\{1, \dots, n_1\} \times \{1, \dots, n_2\}$ and whose lines are all sets of the form $\{i\} \times \{1, \dots, n_2\}$ for $i \in \{1, \dots, n_1\}$ and those of the form $\{1, \dots, n_1\} \times \{i\}$ for $i \in \{1, \dots, n_2\}$. If $n_1 = n_2$, then we call the grid *symmetrical* and otherwise it is *nonsymmetrical*. The point-line dual of a symmetrical/nonsymmetrical grid is called a symmetrical/nonsymmetrical *dual grid*.

Proposition 1.2.2 ([63, Theorem 5.16]). *If \mathcal{S} is a finite generalized quadrangle in which every point is incident with precisely two lines, then \mathcal{S} is a grid. Dually, if \mathcal{S} is a finite generalized quadrangle in which every line is incident with precisely two points, then \mathcal{S} is a dual grid.*

Proposition 1.2.3 ([63, Theorem 5.17]). *Let \mathcal{S} be a finite generalized quadrangle. Then precisely one of the following holds.*

- (1) *\mathcal{S} is a nonsymmetrical grid.*

- (2) \mathcal{S} is a nonsymmetrical dual grid.
- (3) \mathcal{S} has an order (s, t) .

In finite generalized polygons with an order, we can do elementary counting to obtain the total number of points and lines. This information is collected in the following result.

Proposition 1.2.4 ([138, Lemma 1.5.4]). *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a finite generalized n -gon of order (s, t) with $n \geq 3$ and let x be a point of \mathcal{S} . For $i \in \{0, \dots, n\}$, let $\mathcal{S}_i(x)$ denote the set of elements of $\mathcal{P} \cup \mathcal{L}$ that are at distance i from x in the incidence graph of \mathcal{S} . Then*

- (1) *for $1 \leq i \leq n-1$, we have $|\mathcal{S}_i(x)| = (t+1)s^{\lfloor i/2 \rfloor}t^{\lfloor (i-1)/2 \rfloor}$, and $|\mathcal{S}_n(x)| = s^{n/2}t^{n/2-1}$ for n even while $|\mathcal{S}_n(x)| = s^{n-1}$ for n odd;*
- (2) *if n is even, then*

$$|\mathcal{P}| = (s+1)\frac{(st)^{\frac{n}{2}} - 1}{st - 1} \text{ and } |\mathcal{L}| = (t+1)\frac{(st)^{\frac{n}{2}} - 1}{st - 1};$$

- (3) *if n is odd, then $s = t$ and*

$$|\mathcal{P}| = |\mathcal{L}| = \frac{s^n - 1}{s - 1}.$$

Since subgeometries of a generalized polygon are going to play an important role in this thesis, we include the following results.

Lemma 1.2.5. *Let \mathcal{S} be a generalized n -gon contained in another generalized n -gon \mathcal{S}' as a full subgeometry.³ Then \mathcal{S} is isometrically embedded in \mathcal{S}' .*

Proof. If there are two points x, y in \mathcal{S} such that $\lfloor n/2 \rfloor \geq d_{\mathcal{S}}(x, y) > d_{\mathcal{S}'}(x, y)$, then we will get a subgeometry of \mathcal{S}' isomorphic to an ordinary m -gon for some $m < n$. \square

Two points/lines of a generalized $2n$ -gon are called *opposite* if they lie at distance $2n$ from each other in the incidence graph. By Proposition 1.1.1, two points of a generalized $2n$ -gon are opposite if and only if they are at distance n from each other in the collinearity graph, and two lines are opposite if and only if they are at distance $n-1$ from each other in the collinearity graph. Proposition 1.2.1(1) implies that opposite lines in a generalized $2n$ -gon have the same number of points on them and that opposite points have the same number of lines through them. Note that for every point x and every line L of a generalized $2n$ -gon, there exists a unique point $\pi_L(x)$ on L which is nearest to x , i.e., for every point y on L with $y \neq \pi_L(x)$ we have $d(x, y) = d(x, \pi_L(x)) + 1$.

Lemma 1.2.6. *Let \mathcal{S} be a generalized $2n$ -gon that is contained in a generalized $2n$ -gon \mathcal{S}' as a full subgeometry. Then:*

- (1) *every point of \mathcal{S}' is opposite to a point of \mathcal{S} ;*
- (2) *every line of \mathcal{S}' is opposite to a line of \mathcal{S} ;*
- (3) *if \mathcal{S} is thick, then \mathcal{S}' is thick (and hence has an order).*

³The condition of being a full subgeometry is not necessary here, but we include it because all our subgeometries will be full.

- Proof.* (1) Let x be a point of \mathcal{S}' and let y be a point of \mathcal{S} at distance $i := d(x, \mathcal{S})$ from x (by Lemma 1.2.5 we can simply denote the distance functions $d_{\mathcal{S}'}$ and $d_{\mathcal{S}}$ by d). Let z be a point of \mathcal{S} which is at distance $n - i$ from y . Then there cannot exist any path of length less than n between x and z as otherwise we will get an ordinary m -gon as a subgeometry for some $m < n$, which is not possible since \mathcal{S}' is a generalized n -gon. Therefore, $d(x, z) = n$, i.e., x and z are opposite.
- (2) Let L be a line of \mathcal{S}' and x an arbitrary point on L . By (1), there exists a point z in \mathcal{S} which is opposite to x . Let $\pi_L(z)$ be the unique point on L nearest to z . Then $d(z, \pi_L(x)) = n - 1$. There exists a unique line through z which contains a point at distance $n - 2$ from $\pi_L(z)$. Let L' be any other line through z which is contained in \mathcal{S} . Then we must have $d(L', L) = n - 1$, i.e., L and L' are opposite.
- (3) This now follows from Proposition 1.2.1(1). □

1.2.1. Important Examples

For every field \mathbb{F} , we have the Desarguesian projective plane $\text{PG}(2, \mathbb{F})$ constructed from the vector space \mathbb{F}^3 by taking the 1-dimensional subspaces as points and 2-dimensional subspaces as lines, with incidence as set containment. This is a generalized triangle. When \mathbb{F} is the finite field of order q , for some prime power q , then we get a finite generalized triangle of order q which has $1 + q + q^2$ points and equally many lines.

Every nonsingular quadric of Witt index 2 gives rise to a generalized quadrangle if we take the points to be the 1-dimensional subspaces contained in the quadric, and lines to be the 2-dimensional subspaces contained in the quadric. In fact, every rank 2 polar space is a generalized quadrangle (see [63, Chapter 7] or [12, Chapter 4] for an introduction to polar spaces). For example, the generalized quadrangle given in Figure 1.1 can be obtained from the symplectic polarity of $\text{PG}(3, 2)$ defined by the alternating bilinear form $f((x_0, x_1, x_2, x_3), (y_0, y_1, y_2, y_3)) = x_0y_1 - x_1y_0 + x_2y_3 - x_3y_2$ (the map $U \mapsto \{x \in \text{PG}(3, 2) \mid f(x, u) = 0 \ \forall u \in U\}$ gives the polarity), and it is denoted by $W(2)$. For more on generalized quadrangles see the standard reference [116].

There is a natural doubling construction which allows us to construct generalized $2n$ -gons from generalized n -gons. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \text{I})$ be a point-line geometry. Then the *double* of \mathcal{S} is the point-line geometry $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \text{I}')$ where $\mathcal{P}' = \mathcal{P} \cup \mathcal{L}$, \mathcal{L}' is the set of all sets $\{x, L\}$ where $(x, L) \in \text{I}$ and I' is set containment.⁴ It can be easily shown that if \mathcal{S} is a generalized n -gon, then its double is a generalized $2n$ -gon in which every line is incident with precisely two points. In fact, the converse also holds. Every generalized $2n$ -gon with two points on each line is the double of a generalized n -gon. Therefore, we get a generalized hexagon H_π by doubling a projective plane π . If π is isomorphic to $\text{PG}(2, q)$, then we denote H_π by $\text{H}(1, q)$, noting that this generalized hexagon has order $(1, q)$. The dual of $\text{H}(1, q)$ will be denoted by $\text{H}(q, 1)$; thus the points of $\text{H}(q, 1)$ are the edges of the incidence graph of $\text{PG}(2, q)$ (the flags of the projective plane), and the lines are the vertices of the incidence graph. The automorphism group of $\text{H}(q, 1)$ is isomorphic

⁴This is basically the incidence graph of \mathcal{S} treated as a point-line geometry in itself.

to a semi-direct product $\text{PTL}_3(q) \rtimes C_2$, and thus has size $2r(q^3 - 1)(q^3 - q)(q^3 - q^2)/(q - 1)$ when q is the r -th power of a prime.

Let L be a 2-dimensional subspace of \mathbb{F}_q^{n+1} (or equivalently a line of the n -dimensional projective space $\text{PG}(n, q)$), where q is a prime power. Let $x = (x_0, \dots, x_n)$ and $y = (y_0, \dots, y_n)$ form a basis of L . Then the Grassmann coordinates of L are $p_{ij} = x_i y_j - x_j y_i$ for $0 \leq i < j \leq n$. Note that these coordinates are independent of the choice of basis of L , up to a scalar factor. Thus, the Grassman coordinates map a line of $\text{PG}(n, q)$ to a point of $\text{PG}(\binom{n+1}{2} - 1, q)$. The split Cayley hexagon $H(q)$ is a generalized hexagon of order (q, q) , when q is a prime power, which can be defined as follows [136], [138, Section 2.4.13]. Consider the quadratic form $Q : \mathbb{F}_q^7 \rightarrow \mathbb{F}_q$ defined as $Q(x) = x_0 x_4 + x_1 x_5 + x_2 x_6 - x_3^2$.

- (a) The points of $H(q)$ are all 1-dimensional subspaces of \mathbb{F}_q^7 which vanish on Q .
- (b) The lines of $H(q)$ are all 2-dimensional subspaces of \mathbb{F}_q^7 which vanish on Q , and whose Grassmann coordinates satisfy $p_{12} = p_{34}, p_{54} = p_{32}, p_{20} = p_{35}, p_{65} = p_{30}, p_{01} = p_{36}$ and $p_{46} = p_{31}$.

We will not be using this definition directly in this thesis, but only some of the well-known properties of split Cayley hexagons. The automorphism group of $H(q)$ is isomorphic to the semi-direct product $G_2(q) \rtimes \text{Aut}(\mathbb{F}_q)$ and thus it has size $rq^6(q^6 - 1)(q^2 - 1)$ when q is an r -th power of a prime [138, Proposition 4.6.7]. We also note that by [138, Corollary 3.5.7], the split Cayley hexagon $H(q)$ is isomorphic to its point-line dual $H(q)^D$ if and only if q is a power of 3.

For every prime power q , we also have a generalized hexagon of order (q^3, q) known as the *twisted triality hexagon*, which we denote by $T(q^3, q)$. Its point-line dual is denoted by $T(q, q^3)$. We do not give its definition here and refer the interested reader to [138, Chapter 2]. We have the following well known inclusion of geometries between the classes of generalized hexagons we have discussed so far :

$$H(q, 1) \subset H(q)^D \subset T(q, q^3).$$

Clearly this is a sequence of full isometrically embedded subgeometries.

The doubling construction can also be used to construct generalized octagons of order $(q, 1)$ from generalized quadrangles of order (q, q) . There is another class of generalized octagons known as the Ree-Tits octagons, which have order (q, q^2) , for q an odd power of 2, and are denoted by $\text{GO}(q, q^2)$. For the definition of this geometry see [137]. If $q = 2^{2e+1}$, then $|\text{Aut}(\text{GO}(q, q^2))| = (2e + 1)q^{12}(q^6 + 1)(q^4 - 1)(q^3 + 1)(q - 1)$ (see [138, Proposition 4.6.10]).

1.2.2. Classification of Finite Generalized Polygons

One of the most important results in the theory of generalized polygons is the following theorem due to Feit and G. Higman [80] (also see [138, Theorem 1.7.1]).

Theorem 1.2.7. *Let \mathcal{S} be a finite generalized n -gon, with $n \geq 3$, of order (s, t) . Then one of the following holds:*

- \mathcal{S} is an ordinary n -gon, and thus $s = t = 1$.

- \mathcal{S} is a non-degenerate projective plane of order $s = t \geq 2$.
- \mathcal{S} is a generalized quadrangle and the number $st(1 + st)/(s + t)$ is an integer.
- \mathcal{S} is a generalized hexagon; the number st is a square if $s, t > 1$, and in that case the following numbers are integers for $u = \sqrt{st}$ and $w = s + t$,

$$\frac{u^2(1 + w + u^2)(1 + u + u^2)}{2(w + u)}, \frac{u^2(1 + w + u^2)(1 - u + u^2)}{2(w - u)}.$$

- \mathcal{S} is a generalized octagon; the number $2st$ is a square if $s, t > 1$, and in that case the following numbers are integers for $u = \sqrt{\frac{st}{2}}$ and $w = s + t$,

$$\frac{u^2(1 + w + 2u^2)(1 + 2u^2)(1 + 2u + 2u^2)}{2(w + 2u)}, \frac{u^2(1 + w + 2u^2)(1 + 2u^2)(1 - 2u + 2u^2)}{2(w - 2u)}.$$

- \mathcal{S} is a generalized dodecagon, and either $s = 1$ or $t = 1$.

For thick finite generalized polygons, Proposition 1.2.1 combined with Theorem 1.2.7 tells us that besides the projective planes, there are only generalized quadrangles, hexagons, and octagons. There are further restrictions on the parameters of these geometries due to Higman, Haemers and Roos [88, 90, 91].

Theorem 1.2.8. *Let \mathcal{S} be a finite generalized $2n$ -gon of order (s, t) with $s, t, n \geq 2$. Then the following hold:*

- (1) if $n = 2$, then $s \leq t^2$ and $t \leq s^2$;
- (2) if $n = 3$, then $s \leq t^3$ and $t \leq s^3$;
- (3) if $n = 4$, then $s \leq t^2$ and $t \leq s^2$.

Even with all these restrictions, we are nowhere near a full classification of finite generalized polygons. There are infinitely many parameters which satisfy the conditions above but for which we have no idea about existence or non-existence of a generalized polygon with those parameters (see [63, Section 5.8] for a full list of parameters for which the existence is known). While there are many families of finite generalized quadrangles besides those coming from rank 2 polar spaces that we discussed in Section 1.2.1 (see [116, Chapter 3]), the situation is quite different for generalized hexagons and octagons. The examples mentioned in Section 1.2.1 are the only known thick finite generalized hexagons and octagons. Moreover, even when there exists a generalized n -gon of a given order, we usually do not have a classification of all generalized n -gons of that order, except for some small cases. For example, it is not known whether $H(3)$ is the unique generalized hexagon of order 3.

For $s = 2$, a finite generalized quadrangle of order (s, t) necessarily has $t \in \{1, 2, 4\}$ by the restrictions above, and for all these cases there is a unique generalized quadrangle of the given order. For $s = 3$, we have $t \in \{1, 3, 5, 6, 9\}$. There is no GQ of order $(3, 6)$. For $t \in \{1, 5, 9\}$ there is a unique GQ. For $t = 3$ there are exactly two GQ's, dual to each other. There is a unique GQ of order $(4, t)$ for $t \in \{1, 2, 4\}$, but the uniqueness is not known for orders $(4, 6)$, $(4, 8)$ and $(4, 16)$ (existence of GQ's of order $(4, 11)$ and

(4, 12) is still unknown). See [116, Chapter 6] for the proofs of all these results and the relevant references. Again by the restrictions above, a finite generalized hexagon of order $(2, t)$ must have $t \in \{1, 2, 8\}$. In [53], Cohen and Tits proved that $H(2, 1)$, $H(2)$, $H(2)^D$ and $T(2, 8)$ are the only finite generalized hexagons of order $(2, t)$. For $s = 3$, we have $t \in \{1, 3, 27\}$, and only for $t = 1$ we know that the generalized hexagon is unique. For the case of generalized octagons, we do not even know if the smallest known thick generalized octagon, the Ree-Tits octagon of order $(2, 4)$, is the unique generalized octagon of order $(2, 4)$.

We finally note that the situation is not so hopeless if we assume some extra conditions on the generalized polygon. For example, using the classification of finite simple groups, Buekenhout and Van Maldeghem [40] have classified all finite generalized polygons whose automorphism group acts distance transitively on the set of points. There are several other interesting conditions that one can impose which give rise to some classification results, see for example [119] and [138, Chapters 4,5].

1.3. Near Polygons

Near polygons were introduced by Shult and Yanushka in 1980 [126] for studying the so-called tetrahedrally closed line systems in Euclidean spaces. They now form an important class of point-line geometries and have close connections to distance-regular graphs [34] and polar spaces [42]. The standard reference for near polygons is [58], which along with [63] will be our main reference for this section.

Definition. A near $2n$ -gon, for $n \in \mathbb{N}$, is a partial linear space $\mathcal{S} = (\mathcal{P}, \mathcal{L}, I)$ which satisfies the following axioms:

- NP1** the collinearity graph of \mathcal{S} is connected with diameter n ;
- NP2** for every point $x \in \mathcal{P}$ and every line $L \in \mathcal{L}$, there exists a unique point $\pi_L(x)$ incident with L that attains the minimum distance between x and points incident with L .

A near polygon is a near $2n$ -gon for some $n \in \mathbb{N}$. A near 0-gon is just a point while a near 2-gon is a line. It is easy to see that every generalized $2n$ -gon is also a near $2n$ -gon. In fact, every generalized $2n$ -gon is a special kind of near $2n$ -gon, as the following result shows.

Proposition 1.3.1 ([63, Theorem 5.14]). *For an integer $n \geq 2$, every near $2n$ -gon \mathcal{S} that satisfies the following conditions is a generalized $2n$ -gon.*

- (1) *Every point of \mathcal{S} is incident with at least two lines.*
- (2) *For every pair of points x, y at distance $i \in \{1, 2, \dots, n-1\}$ in the collinearity graph of \mathcal{S} , there exists a unique point of \mathcal{S} at distance $i-1$ from x and distance 1 from y .*

Looking at generalized $2n$ -gons as special kinds of near $2n$ -gons will be quite useful to us, and we will do so whenever convenient. If we remove Condition (1) in Proposition 1.3.1, then we get the so-called *degenerate* generalized polygons; for example, a set of lines through a common point gives rise to a degenerate generalized quadrangle.

A simple class of near polygons which are not generalized polygons is the class of Hamming near $2n$ -gons, $n \geq 3$, defined as follows: for arbitrary nonempty sets L_1, \dots, L_n , take $\mathcal{P} = L_1 \times \dots \times L_n$ and let \mathcal{L} to be the union of sets \mathcal{L}_i 's, for $i \in \{1, \dots, n\}$, where $\mathcal{L}_i = \{\{x_1\} \times \dots \times \{x_{i-1}\} \times L_i \times \{x_{i+1}\} \times \dots \times \{x_n\} \mid x_j \in L_j \text{ for } j \in \{1, \dots, n\} \setminus \{i\}\}$; then $(\mathcal{P}, \mathcal{L})$ with incidence as set containment is a near $2n$ -gon. Another example is the near $2n$ -gon \mathbb{H}_n of order $(2, \binom{n+1}{2} - 1)$ obtained by taking the partitions of $\{1, \dots, 2n+2\}$ into $n+1$ pairs as points and partitions into $n-1$ pairs and a 4-subset as lines, a point being incident with a line if it is a refinement of the line. When n is equal to 2, this gives us the generalized quadrangle of Figure 1.1. For a proof of the fact that \mathbb{H}_n is a near $2n$ -gon, see [58, Section 6.2]. There are also some near polygons related to sporadic finite simple groups. For example, the near hexagon \mathbb{E}_2 of order $(2, 14)$ related to the Mathieu group M_{24} , which is constructed by taking the 759 blocks of the Witt design $S(5, 8, 24)$ as points and triples of pairwise disjoint blocks as lines (see [58, Chapter 6] for more on this near polygon). We will later encounter the Hall-Janko near octagon HJ related to the Hall-Janko group J_2 (see [30, 52] for its definition and properties).

Since a near polygon does not have any ordinary triangles as subgeometries, the maximal cliques of its collinearity graph correspond bijectively to the lines of the near polygon. This leads to the following result on isomorphisms between near polygons.

Lemma 1.3.2 ([63, Theorem 6.2]). *If \mathcal{S}_1 and \mathcal{S}_2 are two near polygons with collinearity graphs Γ_1 and Γ_2 , respectively, then \mathcal{S}_1 is isomorphic to \mathcal{S}_2 if and only if Γ_1 is isomorphic to Γ_2 .*

For near polygons that have an order, we have the following easy but rather useful result.

Lemma 1.3.3 ([58, Theorem 1.2]). *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a finite near polygon of order (s, t) and let x be a point of \mathcal{S} . Then*

$$\sum_{y \in \mathcal{P}} \left(\frac{-1}{s} \right)^{d(x, y)} = 0.$$

Proof. By **NP2**, for every line L the sum $\sum_{y \in L} \left(\frac{-1}{s} \right)^{d(x, y)}$ is 0. Therefore, we have

$$0 = \sum_{L \in \mathcal{L}} \sum_{y \in L} \left(\frac{-1}{s} \right)^{d(x, y)} = \sum_{y \in \mathcal{P}} \sum_{L \in \mathbf{I}_y} \left(\frac{-1}{s} \right)^{d(x, y)} = (t+1) \sum_{y \in \mathcal{P}} \left(\frac{-1}{s} \right)^{d(x, y)}.$$

□

Definition. A regular near $2n$ -gon is a near $2n$ -gon \mathcal{S} with an order (s, t) and constants t_i , for $i \in \{0, \dots, n\}$, such that for every two points x and y of \mathcal{S} at distance i from each other, there are precisely $t_i + 1$ lines through y containing a point at distance $i - 1$ from x . Clearly, $t_0 = -1$, $t_1 = 0$ and $t_n = t$. We say that \mathcal{S} is regular with parameters $(s, t; t_2, \dots, t_{n-1})$.

Lemma 1.3.4 ([58, Theorem 1.25]). *The regular near $2n$ -gons, $n \geq 1$, are precisely those near $2n$ -gons whose collinearity graph is distance regular.*

For the definition and properties of distance regular graphs, we refer to [34]. We note that the Hall-Janko near octagon is a regular near octagon of parameters $(2, 4; 0, 3)$ and

in fact Cohen and Tits proved that it is the unique regular near octagon with these parameters [53]. It is also well-known that the Hall-Janko near octagon contains the generalized hexagon $H(2)^D$ as a subgeometry. Note that generalized $2n$ -gons with an order (s, t) are precisely the regular near $2n$ -gons with parameters $(s, t; 0, \dots, 0)$. Just like the case of generalized polygons, elementary counting gives us the following result for regular near polygons.

Lemma 1.3.5 ([58, Chapter 3]). *Let \mathcal{S} be a near $2n$ -gon with parameters s, t, t_i with $i \in \{0, \dots, n\}$. Let x be a point of \mathcal{S} and Γ the collinearity graph of \mathcal{S} . Then the number $|\Gamma_i(x)|$, for $i \in \{0, \dots, n\}$ is independent from the chosen point x and is equal to*

$$k_i = \frac{s^i \prod_{j=0}^{i-1} (t - t_j)}{\prod_{j=1}^i (t_j + 1)},$$

and the total number of points in \mathcal{S} is equal to $v = k_0 + \dots + k_n$.

We include another characterization of generalized $2n$ -gons in terms of near $2n$ -gons. Recall that the number of points in a generalized $2n$ -gon of order (s, t) is equal to $(1 + s)(1 + st + s^2t^2 + \dots + s^{n-1}t^{n-1})$ (see Proposition 1.2.4).

Lemma 1.3.6. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, I)$ be a near $2n$ -gon, $n \geq 2$, of order (s, t) . Then we have*

$$|\mathcal{P}| \leq (1 + s)(1 + st + s^2t^2 + \dots + s^{n-1}t^{n-1}),$$

with equality if and only if \mathcal{S} is a generalized $2n$ -gon.

Proof. Let x be a point of \mathcal{S} and denote the set of points at distance i from x by $\Gamma_i(x)$. By **NP2**, for every $i \in \{1, \dots, n\}$ there is no line completely contained in $\Gamma_i(x)$. Since there are $t + 1$ lines through x , each containing s points besides x , we have $|\Gamma_1(x)| = s(t + 1)$. Similarly, every point of $\Gamma_1(x)$ is collinear with exactly st points of $\Gamma_2(x)$, and every point of $\Gamma_2(x)$ is collinear with at least one point of $\Gamma_1(x)$. Therefore, we have $|\Gamma_2(x)| \leq st|\Gamma_1(x)| = s^2t(t + 1)$. Similarly for each $i \in \{2, \dots, n - 1\}$ we get $|\Gamma_i(x)| \leq st|\Gamma_{i-1}(x)|$, and $|\Gamma_n(x)| \leq st|\Gamma_{n-1}(x)|/(t + 1)$. From all these inequalities we can see that

$$|\mathcal{P}| \leq 1 + \sum_{i=1}^{n-1} s^i t^{i-1} (t + 1) + s^n t^{n-1} = (1 + s)(1 + st + s^2t^2 + \dots + s^{n-1}t^{n-1}),$$

with equality if and only if equality occurs at each “level”, which is equivalent to \mathcal{S} being a regular near polygon of parameters $(s, t; 0, \dots, 0)$, i.e., a generalized $2n$ -gon. \square

1.3.1. Substructures of a Near Polygon

One of the most fundamental structural results for near polygons is the existence of certain substructures called *quads*, which is going to be crucial in many of our proofs later. This result was proved by Shult and Yanushka in [126] and later generalized by Brouwer and Wilbrink in [38] to include existence of sub near polygons of larger diameter. We mention the main results regarding quads and convex subspaces that we will need and refer the reader to [58] for more details.

A subset X of points in a partial linear space $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ is called a *subspace* if for every pair of distinct collinear points in X , all the points of the line joining these points are contained in X . A subspace is called *convex* if for every pair of points $x, y \in X$, every point on every shortest path between x and y in the collinearity graph is contained in X . It is easy to see that every subset X of points of a partial linear space $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ is contained in a unique smallest convex subspace $C(X)$. For every nonempty convex subspace X of \mathcal{S} , we can define a full isometrically embedded subgeometry of \mathcal{S} by taking the elements of X as the points of the subgeometry and all the lines L whose every point is contained in X as the lines of the subgeometry. We call this subgeometry, the *subgeometry induced by X* . We begin with the simplest non-trivial convex subspaces, the lines.

Lemma 1.3.7 ([38, Lemma 1]). *Let L_1, L_2 be two lines of a near polygon \mathcal{S} . Then one of the following two cases occurs:*

- (1) *There exists a point $y_1 \in L_1$ and a point $y_2 \in L_2$ such that $d(L_1, L_2) = d(y_1, y_2)$ and for every $x_1 \in L_1$ and $x_2 \in L_2$, we have $d(x_1, x_2) = d(x_1, y_1) + d(y_1, y_2) + d(y_2, x_2)$.*
- (2) *For every point $x_1 \in L_1$, there exists a unique point $x_2 \in L_2$ such that $d(x_1, x_2) = d(L_1, L_2)$, and for every point $y_2 \in L_2$, there exists a unique point $y_1 \in L_1$ such that $d(y_1, y_2) = d(L_1, L_2)$. In this case we say that the lines L_1 and L_2 are parallel.⁵*

Definition. A *quad* Q of a near polygon \mathcal{S} is a convex subspace of \mathcal{S} that induces a subgeometry isomorphic to a (non-degenerate) generalized quadrangle.

The following result of Shult and Yanushka on existence of quads in a near polygon will be used implicitly in our arguments later.

Theorem 1.3.8 ([126, Proposition 2.5]). *Let a, b be two points of a near polygon \mathcal{S} at distance 2 from each other with more than one common neighbor. Let c, d be two common neighbors of a and b . If at least one of the lines ac, ad, bc, bd contains at least three points, then a and b are contained in a unique quad of \mathcal{S} .*

Most of the near polygons studied in this thesis will have three points on each line, for which Theorem 1.3.8 says that every two points at distance 2 from each other which have more than one common neighbor have a unique quad through them. This quad is a generalized quadrangle of order $(2, t)$ for some possibly infinite cardinal number t . It was proved by Cameron in [41] that for every generalized quadrangle of order $(2, t)$, the parameter t must be finite. Thus by Theorems 1.2.7 and 1.2.8, we must have $t \in \{1, 2, 4\}$. In view of Theorem 1.3.8, this implies the following useful lemma.

Lemma 1.3.9. *If \mathcal{S} is a (possibly infinite) near polygon with three points on each line, then every two points of \mathcal{S} at distance 2 from each other have either 1, 2, 3 or 5 common neighbors.*

It is not difficult to classify all finite generalized quadrangles of order $(2, t)$ (see for example [58, Section 1.10]). For $t = 1$ we have the (3×3) -grid. The remaining cases, $t = 2$ and $t = 4$ give us $W(2)$ (see Figure 1.1) and $Q(5, 2)$ (see [116, Section 3.1] for its definition), respectively. The corresponding quads of a near polygon with three points on each line will be called grid-quads, $W(2)$ -quads and $Q(5, 2)$ -quads, respectively. We will also need the following property of quads which is a direct consequence of Theorem 1.3.8.

⁵Note that parallelism is not a transitive relation here.

Lemma 1.3.10. *Let L_1, L_2 be two intersecting lines of a near polygon \mathcal{S} . Then there is at most one quad of \mathcal{S} which contains both L_1 and L_2 .*

Let x be a point of a near polygon \mathcal{S} and F a convex subspace of \mathcal{S} (for example, a quad or a line). Then we say that the point x is *classical* with respect to F , if there exists a (necessarily unique) point $\pi_F(x) \in F$ such that $d(x, y) = d(x, \pi_F(x)) + d(\pi_F(x), y)$ for every $y \in F$. The point $\pi_F(x)$ is called the *projection* of x on F . It follows from **NP2**, that every point is classical with respect to a given line. We call a convex subspace F of \mathcal{S} *classical* if every point of \mathcal{S} is classical with respect to F . This notion of classical subspaces will be studied in a much more general setting in Chapter 2. Cameron proved in [42] that every near polygon in which every pair of points at distance 2 have a unique quad through them and every quad is classical must be a dual polar space.

Lemma 1.3.11 ([63, Theorem 6.8]). *Let F be a nonempty convex subspace of a near polygon \mathcal{S} and let x be a point of \mathcal{S} such that $d(x, F) \leq 1$. Then x is classical with respect to F .*

Lemma 1.3.12 ([63, Theorem 6.9]). *The intersection of two classical convex subspaces F_1, F_2 of a near polygon \mathcal{S} is either empty or a classical convex subspace.*

1.3.2. Near Hexagons of Order 2

In this section we apply the basic theory of near polygons mentioned so far, along with some well-known results, to classify all near hexagons of order 2. This result fits better in Chapter 6 where – among other things – we prove that the dual split Cayley hexagon $H(2)^D$ is the unique near hexagon of order $(2, 2)$ which contains the generalized hexagon $H(2, 1)$ as a full isometrically embedded subgeometry. But due to its elementary nature we prove the result here.

Both the split Cayley hexagon $H(2)$ and its dual $H(2)^D$ are examples of near hexagons of order 2. Another example is the Hamming near hexagon $\mathbb{L}_3 \times \mathbb{L}_3 \times \mathbb{L}_3$ where \mathbb{L}_3 is just a set (line) of three elements (points)⁶. We will show that there are no more near hexagons of order 2. Note that in a generalized hexagon of order 2 every pair of points at distance 2 have a unique common neighbor, while in the Hamming near hexagon of order 2 every pair of points at distance 2 from each other have exactly 2 common neighbors.

Lemma 1.3.13. *Let \mathcal{S} be a finite near hexagon of order (s, t) and Q a quad of \mathcal{S} that has order (s, t') . Then $t' < t$.*

Proof. We know that $t' \leq t$. For the sake of contradiction, assume that $t' = t$. Let x be a point of Q . Since $t = t'$, all lines of \mathcal{S} through x are already contained in Q , and thus x cannot be collinear with any point that is not contained in Q . But then, there cannot be any points of \mathcal{S} that lie outside Q , as the collinearity graph of \mathcal{S} is connected. Thus $\mathcal{S} = Q$, which is a contradiction. \square

Lemma 1.3.14. *Let \mathcal{S} be a near hexagon of order 2. Then the number of common neighbors of a pair of points at distance 2 from each other is a constant $k \in \{1, 2\}$.*

⁶Picture it as a Rubik's cube.

Proof. Let v denote the total number of points of \mathcal{S} . For a fixed point x let $n_i(x)$ denote the number of points at distance $i \in \{0, 1, 2, 3\}$ from x (in the collinearity graph). For all $x \in \mathcal{P}$, we have $n_0(x) = 1$, $n_1(x) = 6$, and thus

$$n_2(x) + n_3(x) = v - 7. \quad (1.1)$$

By Lemma 1.3.3 we have

$$n_0(x) - \frac{n_1(x)}{2} + \frac{n_2(x)}{4} - \frac{n_3(x)}{8} = 0. \quad (1.2)$$

Solving equations (1.1) and (1.2) we get that $n_2(x) = (v + 9)/3$ and $n_3(x) = (2v - 30)/3$ for all $x \in \mathcal{P}$. Therefore these numbers only depend on v , and we can define constants

$$n_0 = 1, n_1 = 6, n_2 = (v + 9)/3, n_3 = (2v - 30)/3. \quad (1.3)$$

By Lemma 1.3.13 all quads of \mathcal{N} are grid-quads. For a point x , let $N(x)$ be the number of grid-quads that contain x . Then the number of points at distance 2 from x that are contained in a grid along with x is equal to $4N(x)$ since there is a unique quad through a pair of points at distance 2 which have more than one common neighbor by Theorem 1.3.8. Double counting edges between $\Gamma_1(x)$ and $\Gamma_2(x)$ we get that $2 \cdot 4N(x) + 1 \cdot (n_2 - 4N(x)) = n_1 \cdot 4$, and hence $N(x) = (63 - v)/12$. So, the total number of grid-quads through a point is a constant given by $N := (63 - v)/12$.

By Lemma 1.3.10, we must have $N \in \{0, 1, 2, 3\}$ as there are only 3 lines through a point. Since $v = 63 - 12N$, using double counting we get that the total number of grid-quads in \mathcal{N} is

$$N(63 - 12N)/9.$$

This number is not an integer if $N \in \{1, 2\}$. Therefore, N must be 0 or 3. If N is 0, then there are no quads, and hence every two points at distance 2 from each other have a unique common neighbor. Say N is equal to 3 and let x, y be a pair of points at distance 2 from each other. Every line through x is contained in precisely two of the three grid-quads through x . Thus for every neighbor z of x , the two lines through z that contain a point of $\Gamma_2(x)$ are contained in grids through x . This implies that there is a grid through x and y , i.e., the number of common neighbors between x and y is 2. \square

Now let \mathcal{S} be a near hexagon of order 2 and let $k \in \{1, 2\}$ be the constant which is equal to the number of common neighbors between any pair of points at distance 2 from each other in \mathcal{S} . If $k = 1$, then it follows from Proposition 1.3.1 that \mathcal{S} is a generalized hexagon, and hence by the classification result of Cohen and Tits [53] it is either isomorphic to $H(2)$ or $H(2)^D$. So, suppose that $k = 2$. Then we can either use the classification result of Brouwer et al. [35] (see number (xi) in Theorem 1.1 of the paper) or simply draw some pictures and convince ourselves that \mathcal{S} must be isomorphic to $\mathbb{L}_3 \times \mathbb{L}_3 \times \mathbb{L}_3$.

Remark. The classification of finite near polygons is much harder than the classification of finite generalized polygons. But some success has been achieved for certain classes of near polygons, like the dense near hexagons and octagons with $k \in \{3, 4\}$ points on each line (see [58]).

2. Valuations of Near Polygons

2.1. Introduction

Consider the following problem:

Given a near polygon \mathcal{S} , determine all near polygons which contains an isomorphic copy of \mathcal{S} as a full subgeometry.

One of the most fruitful tools to tackle this problem has been the theory of *valuations*, introduced by De Bruyn and Vandecasteele in [64], where one computes certain integer valued functions on the point set of a near polygon \mathcal{S} satisfying some well chosen axioms and then uses them to determine how near polygons containing \mathcal{S} as a full subgeometry look like. For example, De Bruyn and Vandecasteele used this theory to classify certain classes of dense near polygons, i.e., near polygons in which every two points at distance 2 have more than one common neighbor, by exploiting the fact that dense near $2n$ -gons contain convex subspaces which induce near $2m$ -gons, for every $m \in \{2, \dots, n-1\}$ [65] (also see [58, Chapters 6–9]). Another important example is the use of valuations, satisfying a different set of axioms than before, by De Bruyn in [62] to prove that the Ree-Tits octagon of order $(2, 4)$ is the unique generalized octagon of order $(2, 4)$ containing a suboctagon of order $(2, 1)$. In fact, a similar result of De Medts and Van Maldegehem [67] on uniqueness of $H(3)$ as a generalized hexagon of order 3 containing a subhexagon of order $(3, 1)$ can also be proved using this theory of valuations. For a survey on the various applications of valuations of a near polygon, see [61].

We now define and study the notion of valuations which will be used in this thesis. The basic definitions and results of this chapter will be used in Chapter 4 to prove non-existence of certain semi-finite generalized hexagons and in Chapter 6 to obtain characterizations of certain Suzuki tower near polygons. Almost all of these results on valuations are already implicit in the literature (for example, in [59, Section 2]), but for the ease of the reader we provide proofs.

2.2. Definitions and Properties

Definition. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, I)$ be a near $2n$ -gon. A function $f : \mathcal{P} \rightarrow \mathbb{Z}$ is called a *semi-valuation* of \mathcal{S} if for every line $L \in \mathcal{L}$, there exists a unique point x_L incident with L such that $f(x) = f(x_L) + 1$ for every $x \neq x_L$ incident with L . A *valuation* of \mathcal{S} is a semi-valuation which attains a minimum value of 0. Two valuations f, g of \mathcal{S} are called *isomorphic* if there exists an automorphism θ of \mathcal{S} such that $f(x) = g(\theta(x))$ for all $x \in \mathcal{P}$.

It directly follows from the definition that given a valuation f of a near polygon \mathcal{S} and a point x with f -value 0, the points at distance i from x have f -value at most i . Thus, the values taken by a valuation of a near $2n$ -gon lie in the set $\{0, 1, \dots, n\}$. If f is a valuation of a near $2n$ -gon \mathcal{S} , then we denote by \mathcal{O}_f the set of points of \mathcal{S} which have f -value 0, and by M_f the maximum value that f attains.

Definition. A *hyperplane* of a partial linear space \mathcal{S} is a proper subset H of the set of points of \mathcal{S} with the property that every line of \mathcal{S} is either contained in H or intersects H in a unique point¹. If a hyperplane does not contain any line of \mathcal{S} , then it is called a *1-ovoid* of \mathcal{S} .

For a valuation f of a near polygon $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$, let H_f be the set $\{x \in \mathcal{P} \mid f(x) < M_f\}$. For every line L , there exists a point x_L on L such that $f(x_L) = f(x) - 1$ for all $x \in L \setminus \{x_L\}$ and thus $f(x_L) \leq M_f - 1 < M_f$. Therefore, the point x_L belongs to H_f . Now say there are two points x and y on L which are contained in H_f . Without loss of generality, let y be distinct from the unique point x_L . Then every point of L has f -value at most $f(y)$, which is less than M_f since y is in H_f . Thus all the points of the line L are contained in H_f . We have just proved that for every valuation f of \mathcal{S} (in fact, one can take a semi-valuation), there is a corresponding hyperplane H_f of \mathcal{S} consisting of those points of \mathcal{S} that have non-maximal f -value. This correspondence between valuations and hyperplanes of \mathcal{S} will be crucial in many of our results. We note that it is possible for two distinct valuations to give rise to the same hyperplane. A hyperplane H will be called of *valuation type* if there exists a valuation f such that $H = H_f$. We now give some standard examples of valuations.

Example. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near $2n$ -gon, for some positive integer $n \geq 2$, such that $\Gamma_n(x) \neq \emptyset$ for every point x .

- (1) Let x be a fixed point of \mathcal{S} . Define $f_x(y) = d(x, y)$ for all $y \in \mathcal{P}$. Then the function f_x is a valuation of \mathcal{S} , which is known as a *classical valuation* of \mathcal{S} . Its hyperplane H_{f_x} is the set of points at distance at most $n - 1$ from x , and such a hyperplane will be called the *singular hyperplane* with *center* x .
- (2) Let O be a 1-ovoid of \mathcal{S} . Define a function $f_O : \mathcal{P} \rightarrow \mathbb{Z}$ by $f_O(x) = 0$ if $x \in O$ and $f_O(x) = 1$ if $x \notin O$. Then f_O is a valuation of \mathcal{S} , which is known as an *ovoidal valuation*, and its associated hyperplane is equal to O .
- (3) Suppose $n \geq 3$. Let x be a fixed point of \mathcal{S} and let O' be a 1-ovoid of the subgeometry of \mathcal{S} induced on the set $\Gamma_n(x)$ of points at distance n from x . Define $f_{x,O'}(y) := d(x, y)$ for all points y at distance at most $n - 1$ from x , $f_{x,O'}(y) = n - 2$ for $y \in O'$ and $f_{x,O'}(y) = n - 1$ for all the remaining points of \mathcal{S} . Then $f_{x,O'}$ is a valuation of \mathcal{S} , which is known as a *semi-classical valuation*. The associated hyperplane of $f_{x,O'}$ is equal to $O' \cup \Gamma_{\leq n-1}(x)$, and this will be called a *semi-singular hyperplane* with *center* x .

We will now show that in a generalized quadrangle which is not a dual grid (recall the definition from Section 1.2), every valuation is either classical or ovoidal. This will later

¹This definition is motivated by the same property of hyperplanes in projective spaces, and in fact when we embed a partial linear space in a projective space as a full subgeometry, the hyperplanes of the projective space give rise to the hyperplanes of the partial linear space [63, Section 4.5].

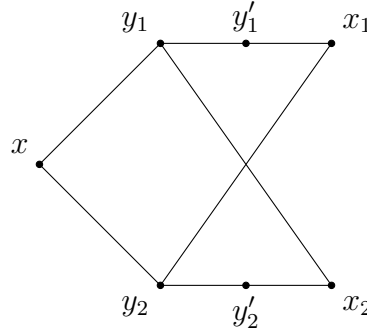


Figure 2.1.: Lemma 2.2.1

be used to derive the point-quad relationships in near polygons, as originally obtained by Shult and Yanushka in [126, Proposition 2.6]².

Lemma 2.2.1. *Let \mathcal{S} be a generalized quadrangle which is not a dual grid and let x be a point of \mathcal{S} . Then the subgraph of the collinearity graph Γ induced on the set $\Gamma_2(x)$ of points at distance 2 from x is connected and has diameter at most 3.*

Proof. If \mathcal{S} is a grid then $\Gamma_2(x)$ is either a subline, or another grid³; and thus the graph has diameter at most 2. So assume that \mathcal{S} is neither a grid, nor a dual grid. Then by Proposition 1.2.3, \mathcal{S} has an order (s, t) . If $t = 1$ or $s = 1$, then \mathcal{S} is a grid or a dual grid. Therefore $s, t \geq 2$, and in particular every line of \mathcal{S} has at least three points.

Now let x_1, x_2 be two non-collinear points of $\Gamma_2(x)$. Then $d(x_1, x_2) = 2$ in \mathcal{S} . Say, there exists a common neighbor of x_1 and x_2 which also lies in $\Gamma_2(x)$. Then x_1, x_2 have distance 2 in the graph induced on $\Gamma_2(x)$. Therefore, assume that all common neighbors of x_1 and x_2 lie in $\Gamma_1(x)$. Let y_1, y_2 be two such common neighbors. Since every line has at least three points, we can pick a point y'_1 on the line x_1y_1 which is not equal to x_1 or y_1 . Then y'_1 must lie in $\Gamma_2(x)$. Clearly y'_1 is not collinear with either x_2 or y_2 as that would contradict the axiom **GQ2** (see Figure 2.1). Let y'_2 be the unique point of the line x_2y_2 which is collinear with y'_1 . Then y'_2 is also contained in $\Gamma_2(x)$, which gives us the path x_1, y'_1, y'_2, x_2 between x_1 and x_2 of length 3. Therefore, the graph is connected with diameter at most 3. \square

Theorem 2.2.2. *Let \mathcal{S} be a generalized quadrangle which is not a dual grid. Then every valuation of \mathcal{S} is either classical or ovoidal.*

Proof. Let f be a valuation of \mathcal{S} . If $M_f = 1$, then every line contains a unique point of value 0, and hence \mathcal{O}_f is a 1-ovoid and f is ovoidal. So, assume that $M_f = 2$. We will show that f is classical. Let $x \in \mathcal{O}_f$ and let y be a point with f -value 2. Then every point in $\Gamma_1(x)$ has f -value 1. Let y' be a point of $\Gamma_2(x)$ which is collinear with y , and let y'' be the unique point on the line yy' which is collinear with x . Then $f(y'') = 1$, $f(y) = 2$, and therefore $f(y')$ must be equal to 2 by the property of valuations. Now by Lemma 2.2.1, every point in $\Gamma_2(x)$ has f -value 2. This shows that f is a classical valuation with x as the unique point with f -value 0. \square

²The proof using valuations is essentially the same as the proof of Shult and Yanushka.

³The case where $\Gamma_2(x)$ is a singleton is not possible as then \mathcal{S} must be a 2×2 grid, which contradicts the fact that \mathcal{S} is not a dual grid since a 2×2 grid is also a dual grid.

Recall the main problem stated in Section 2.1. We will now see why valuations can be useful in tackling that problem.

Theorem 2.2.3. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a near polygon which is a full subgeometry of another near polygon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathcal{I}')$ ⁴. Then for every point x in \mathcal{P}' , the function $f_x : \mathcal{P} \rightarrow \mathbb{Z}$ defined by $f_x(y) = d(x, y) - d(x, \mathcal{P})$ is a valuation of \mathcal{S} . Moreover, if \mathcal{S} is isometrically embedded, then for every $x \in \mathcal{P}$ the valuation f_x is a classical valuation of \mathcal{S} .*

Proof. Let $L \in \mathcal{L}$ be a line of \mathcal{S} , and let $x_L = \pi_L(x)$ be the unique point on L which is nearest to x in \mathcal{S}' . Since \mathcal{S} is a full subgeometry, $x_L \in \mathcal{P}$. Then for every point y on L , we have $d(x, y) = d(x, x_L) + 1$, and hence x_L is the unique point on L with minimum f_x -value. Taking a point $y \in \mathcal{P}$ with $d(x, y) = d(x, \mathcal{P})$, we see that f_x attains the value 0. Thus, f_x is a valuation of \mathcal{S} .

Now assume that \mathcal{S} is isometrically embedded and let $x \in \mathcal{P}$. Then for every $y \in \mathcal{P}$ we have $f_x(y) = d(x, y) = d_{\mathcal{S}}(x, y)$, and hence f_x is classical with center x . \square

In view of Theorem 2.2.3, whenever we have a near polygon $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ as a full subgeometry of another near polygon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathcal{I}')$ and x is a point of \mathcal{S}' , then f_x will denote the valuation of \mathcal{S} defined by $f_x(y) = d(x, y) - d(x, \mathcal{P})$ for $y \in \mathcal{P}$. This will sometimes be referred as the valuation of \mathcal{S} corresponding to the point x of \mathcal{S}' .

Lemma 2.2.4. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ be a near polygon which is a full subgeometry of a near $2n$ -gon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathcal{I}')$. For every $x \in \mathcal{P}'$, let f_x be the corresponding valuation of \mathcal{S} . Then we have the following:*

- (1) *If $d(x, \mathcal{P}) \geq i$, then $M_{f_x} \leq n - i$.*
- (2) *If $d(x, \mathcal{P}) = 1$ and $|\mathcal{O}_{f_x}| = 1$, then x is collinear with a unique point of \mathcal{S} .*

Proof. For a point $x \in \mathcal{P}'$, the maximum possible distance between x and a point of \mathcal{P} is n . Since $f_x(y) = d(x, y) - d(x, \mathcal{P}) \leq n - i$ for every $y \in \mathcal{P}$, we have $M_{f_x} \leq n - i$. If $d(x, \mathcal{P}) = 1$, then the set \mathcal{O}_{f_x} corresponds to the points of \mathcal{S} which are at distance 1 from x . \square

Theorem 2.2.3 shows that if we are given a near polygon \mathcal{S} , and if we can somehow determine the set V of all possible valuations of \mathcal{S} , then given any near polygon \mathcal{S}' containing \mathcal{S} as a full subgeometry, to each point of \mathcal{S}' we can associate an element of V . This mapping is usually neither injective, nor surjective. But in some cases — as we will see later — we can reconstruct \mathcal{S}' using some well chosen elements of V (which we have already computed). But first, we derive the well-known point-quad relations in a near polygon. Given a quad Q of a near polygon \mathcal{S} , recall that we call a point x classical with respect to Q if there exists a point $\pi_Q(x) \in Q$ such that for every $y \in Q$ we have $d(x, y) = d(x, \pi_Q(x)) + d(\pi_Q(x), y)$. We call a point ovoidal with respect to Q if the set of points in Q which are at distance $d(x, Q)$ from x , i.e., points of Q which are nearest to x , form a 1-ovoid of Q .

Theorem 2.2.5. *Let x be a point and Q a quad of a near polygon \mathcal{S} such that Q is not a dual grid. Then either x is classical with respect to Q or x is ovoidal with respect to Q .*

Proof. This is an immediate corollary of Theorems 2.2.2 and 2.2.3 as $d(x, y) = f_x(y) + d(x, Q)$ for every $y \in Q$. \square

⁴Note that we do not make the assumption that these near polygons have the same diameter.

For a near polygon \mathcal{S} , we have associated points of an arbitrary near polygon \mathcal{S}' containing \mathcal{S} as a full subgeometry to elements of the set V of all valuations of \mathcal{S} . Of course, it will help if we can associate lines of \mathcal{S}' to certain subsets of V . This motivates the following definition.

Definition. Let f, g be two valuations of a near polygon \mathcal{S} . Then we say that f and g are *neighboring* valuations if there exists an integer ϵ such that $|f(x) - g(x) + \epsilon| \leq 1$ for every point x of \mathcal{S} .

Note that the integer ϵ defined above must belong to the set $\{-1, 0, 1\}$. This can be shown as follows. If $\epsilon \leq -2$, then for a point x in \mathcal{O}_f we have $f(x) - g(x) + \epsilon = 0 - g(x) + \epsilon \leq -2$ since $g(x) \geq 0$, which is a contradiction to the fact that $|0 - g(x) + \epsilon| \leq 1$. Therefore $\epsilon \geq -1$. Similarly, if $\epsilon \geq 2$, then for a point $y \in \mathcal{O}_g$ we have $f(y) - g(y) + \epsilon = f(y) - 0 + \epsilon \geq 2$ since $f(y) \geq 0$, which is again a contradiction as $|f(y) - 0 + \epsilon| \leq 1$. Therefore, $\epsilon \leq 1$. Moreover, the number ϵ is uniquely determined by f and g unless $f = g$, in which case ϵ can be taken to be any number in $\{-1, 0, 1\}$ (see for example [59, Corollary 2.3]).

Theorem 2.2.6. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near polygon which is a full subgeometry of another near polygon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathbf{I}')$. Then for every line L of \mathcal{L}' , $\{f_x \mid x \in L\}$ is a set of pairwise neighboring valuations of \mathcal{S} .

Proof. Let L be a line of \mathcal{S}' , and x, y two points on L . Then for every point $z \in \mathcal{P}$, we have $|d(z, x) - d(z, y)| \leq 1$ by the triangle inequality. Now write $d(z, x) = f_x(z) + d(x, \mathcal{P})$, $d(z, y) = f_y(z) + d(y, \mathcal{P})$ and $\epsilon = d(x, \mathcal{P}) - d(y, \mathcal{P})$. Then for every $z \in \mathcal{P}$, we have $|f_x(z) - f_y(z) + \epsilon| \leq 1$, which shows that f_x and f_y are neighboring valuations. \square

2.3. Three Points on Each Line

We now restrict to the case when the near polygons have three points on each line, i.e., we have a near polygon \mathcal{S} contained in another near polygon \mathcal{S}' as a full subgeometry and every line of \mathcal{S}' is incident with exactly three points. Then by applying Theorems 2.2.3 and 2.2.6, to the points of \mathcal{S}' we can associate valuations of \mathcal{S} and to the lines of \mathcal{S}' we can associate triples of pairwise neighboring valuations of \mathcal{S} . We will show that it suffices to focus on triples of distinct valuations and see how all the information regarding valuations can be written in a “compact form”. All the results of this section are essentially contained in [59]. But there the notation used is different and the results proved are slightly more general. Therefore, it is worthwhile to revisit the proofs.

Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near polygon with three points on each line. Let f_1 and f_2 be two neighboring valuations of \mathcal{S} and let $\epsilon \in \{-1, 0, 1\}$ be such that $|f_1(x) - f_2(x) + \epsilon| \leq 1$ for all $x \in \mathcal{P}$. We define a function $f'_3 : \mathcal{P} \rightarrow \mathbb{Z}$ as follows. For a point $x \in \mathcal{P}$, if $f_1(x) = f_2(x) - \epsilon$, then we define $f'_3(x) := f_1(x) - 1 = f_2(x) - \epsilon - 1$. For every other point x , we define $f'_3(x) := \max\{f_1(x), f_2(x) - \epsilon\}$. Let m (necessarily belonging to $\{-1, 0, 1\}$) be the minimum value taken by f'_3 , and define $f_3 : \mathcal{P} \rightarrow \mathbb{N}$ as $f_3(x) := f'_3(x) - m$ (so that f_3 attains the value 0). Then we denote this function f_3 by $f_1 * f_2$. Note that this function is well defined in the sense that if $f_1 = f_2$, then for any of the three possible values of ϵ , we have $f_1 * f_2 = f_1 = f_2$. It also follows directly from the definition that $f_1 * f_2 = f_2 * f_1$.

Lemma 2.3.1. Let f_1 and f_2 be two neighboring valuations of a near polygon \mathcal{S} that has three points on each line. Let $f_3 = f_1 * f_2$. Then

- (1) f_3 is also a valuation of \mathcal{S} ;
- (2) f_1, f_2 and f_3 are all pairwise neighboring;
- (3) we have $f_2 * f_3 = f_1$ and $f_1 * f_3 = f_2$.

Proof. (1) It suffices to show that the function f'_3 defined above is a semi-valuation. We show this under the weaker assumption that f_1 and f_2 are semi-valuations, and thus we can simply assume that $|f_1(x) - f_2(x)| \leq 1$ for all $x \in \mathcal{P}$ (every translate of a valuation is a semi-valuation and there is a unique translate of a semi-valuation which is a valuation).

Let L be a line of \mathcal{S} . For $i \in \{1, 2\}$, let x_i be the unique point on L such that $f_i(x) = f_i(x_i) + 1$ for all $x \in L \setminus \{x_i\}$. Say $x_1 = x_2$. Then for all $x \in L \setminus \{x_1\}$, we have $f_2(x) = f_2(x_2) + 1 = f_1(x_1) + c + 1 = f_1(x) + c$, where $c = f_2(x_1) - f_1(x_1)$. If $c = 0$, then $f'_3 = f_1 - 1$, and otherwise f_3 is either equal to f_1 or f_2 depending on the sign of c . Thus, in all these cases f'_3 is a semi-valuation.

Now say $x_1 \neq x_2$. If $f_1(x_1) > f_2(x_2)$, then $f_1(x_2) - f_2(x_2) = f_1(x_1) + 1 - f_2(x_1) \geq 2$, a contradiction. Similarly, we cannot have $f_1(x_1) < f_2(x_2)$ and therefore $f_1(x_1) = f_2(x_2) = a$. Let x_3 be the third point of L . Then $f'_3(x_1) = f'_3(x_2) = a + 1$ and $f'_3(x_3) = a$. Thus, f'_3 is a semi-valuation.

- (2) We show that f_1 and f_3 are neighboring and the other cases follow from symmetry. Let x be a point for which $f'_3(x) = f_1(x) - 1$. Since $f_3(x) = f'_3(x) - m$, where m is the minimum value of f'_3 , we get $f_1(x) - f_3(x) + (-m) = 1$. Now let x be any other point. Say $f_1(x) \geq f_2(x) - \epsilon$. Then $f_1(x) - f_3(x) + (-m) = 0$. Otherwise, $f_2(x) - \epsilon = f_3(x) + m$, but since $|f_1(x) - f_2(x) + \epsilon| \leq 1$, we get $|f_1(x) - f_3(x) + (-m)| \leq 1$. This shows that there exists an integer $-m$ such that $|f_1(x) - f_3(x) + (-m)| \leq 1$ for all x , which means that f_1 and f_3 are neighboring.
- (3) Again, we will only prove that $f_1 * f_3 = f_2$. We have just seen that f_1 and f_3 are neighboring with $|f_1(x) - f_3(x) + (-m)| \leq 1$ for all $x \in \mathcal{P}$, where m is the minimum value taken by the function f'_3 defined before. Define f' as $f'(x) = f_1(x) - 1$ whenever $f_1(x) = f_3(x) - (-m)$ and $f'(x) = \max\{f_1(x), f_3(x) - (-m)\}$ otherwise. Note that $f_3(x) - (-m) = f'_3(x)$. Let x be a point where $f_1(x) = f'_3(x)$. This only holds for $f_1(x) > f_2(x) - \epsilon$ and then $f'(x) = f_1(x) - 1 = f'_3(x) - 1$. Since $|f_1(x) - f_2(x) + \epsilon| < 1$, we must have $f_1(x) - 1 = f'_3(x) - 1 = f_2(x) - \epsilon$; thus $f'(x) = f_2(x) - \epsilon$.

Now let x be a point where $f_1(x) \neq f'_3(x)$. Then we must have $f_2(x) - \epsilon \geq f_1(x)$. If there is equality, then $f'_3(x) = f_1(x) - 1$ and thus $f'(x) = \max\{f_1(x), f'_3(x)\} = f_1(x) = f_2(x) - \epsilon$. If not, then $f'_3(x) = f_2(x) - \epsilon > f_1(x)$, and hence $f'(x) = f_2(x) - \epsilon$. Therefore, f' is just a translate of the valuation f_2 by $-\epsilon$, which implies that $f_1 * f_3 = f_2$.

□

The main addition to Theorems 2.2.3 and 2.2.6 for the case of three points per line is the following result, which leads us to the definition of *Valuation Geometry* of a near polygon with three points on each line.

Theorem 2.3.2. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near polygon that is a full subgeometry of another near polygon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathbf{I}')$ and suppose every line in \mathcal{S}' is incident with precisely three points. Let $\{x_1, x_2, x_3\}$ be a line of \mathcal{S}' and let f_1, f_2, f_3 be the valuations of \mathcal{S} induced by*

the points x_1, x_2, x_3 , respectively. Then (a) $f_3 = f_1 * f_2$; (b) if every point of the line is at the same distance from \mathcal{P} , then all three valuations are mutually distinct.

Proof. For each $i \in \{1, 2, 3\}$, let $d_i = d(x_i, \mathcal{P})$. We have seen that f_1 and f_2 are neighboring valuations with $\epsilon = d_1 - d_2$. Define f'_3 as before. We will show that f_3 is a translate of f'_3 , and since there is only one possible translate of a semi-valuation which is a valuation this will show that $f_3 = f_1 * f_2$.

Let x be a point of \mathcal{S} . Say $d(x, x_1) = d(x, x_2)$. This is equivalent to $f_1(x) = f_2(x) + d_2 - d_1 = f_2(x) - \epsilon$. Then by **NP2**, we have $d(x, x_3) = d(x, x_1) - 1$, which in terms of valuations means $f_3(x) + d_3 = f_1(x) + d_1 - 1$. Therefore, $f_3(x) = f_1(x) - 1 + \epsilon'$, where $\epsilon' = d_1 - d_3$. Now say $d(x, x_1) \neq d(x, x_2)$. Then again by **NP2**, $d(x, x_3) = \max\{d(x, x_1), d(x, x_2)\}$. In terms of valuations, we have $f_3(x) + d_3 = \max\{f_1(x) + d_1, f_2(x) + d_2\} = \max\{f_1(x) + d_1, f_2(x) + d_2 - d_1 + d_1\}$, which is equivalent to $f_3(x) = \max\{f_1(x) + \epsilon', f_2(x) - \epsilon + \epsilon'\} = \max\{f_1(x), f_2(x) - \epsilon\} + \epsilon'$. Therefore, $f_3(x) = f'_3(x) + \epsilon'$ for all $x \in \mathcal{P}$ and hence $f_3 = f_1 * f_2$.

This shows that if any two of these valuations coincide, then all of them coincide. Therefore, it suffices to prove that we cannot have all of the valuations equal if $d_1 = d_2 = d_3$. Say $f := f_1 = f_2 = f_3$ and look at a point y in \mathcal{O}_f . Then we have $d(x_i, y) = f_i(y) + d_i$ for each i , which implies that $d(x_1, y) = d(x_2, y) = d(x_3, y)$. This contradicts **NP2**. \square

Definition. Let \mathcal{S} be a near polygon with three points on each line and let V be the set of all valuations of \mathcal{S} . Then the *valuation geometry* \mathcal{V} of \mathcal{S} is the partial linear space defined by taking the set V as the point set and the set of 3-element subsets $\{f_1, f_2, f_3\}$ of pairwise distinct and neighboring valuations which satisfy $f_3 = f_1 * f_2$ as the line set.

Note that the valuation geometry of \mathcal{S} does not directly give us any information about those lines of a near polygon \mathcal{S}' containing \mathcal{S} as a full subgeometry for which all three points on the line induce the same valuation. But in all of the cases considered in this thesis, we will show that no two collinear points of \mathcal{S}' can induce the same valuation of \mathcal{S} , from which it will follow directly that each line of \mathcal{S}' induces a line of the valuation geometry of \mathcal{S} .

We now see an example of how this valuation geometry can be useful, which will also be used to set up the notation for valuation geometries of near polygons. As a by product, we will get a proof of the uniqueness of the generalized quadrangle of order 2 using valuations. While this proof is in no way simpler than other proofs of uniqueness, it sets the template for more complicated arguments involving valuations that we will see in Chapters 4 and 6.

2.3.1. An Example – the Doily

For this section let \mathcal{S} be the (3×3) -grid, i.e., the unique generalized quadrangle of order $(2, 1)$.⁵ Then by Theorem 2.2.2, every valuation of \mathcal{S} is either classical or ovoidal. From the transitivity of the automorphism group on the point set of \mathcal{S} , it follows that all classical valuations of \mathcal{S} are isomorphic to each other. It is also easy to show that all ovoidal valuations of \mathcal{S} are isomorphic. Let's call the members of the isomorphism class of classical valuations as “type A” valuations, and the ovoidal ones as “type B” valuations

⁵Its uniqueness is quite easy to show using the GQ axioms.

(in general we will be giving labels to elements of an equivalence class of the valuation isomorphism). To every line of the valuation geometry \mathcal{V} of \mathcal{S} , assign a type based on the types of the valuations it contains; which in this case is just a sorted string of A's and B's. For example, if in a line $\{f_1, f_2, f_3\}$, f_1 is classical and f_2, f_3 are ovoidal, then we will call the line to be of "type ABB ". Then the valuation geometry \mathcal{V} , which can either be computed by hand or by the computer algorithms of Chapter 3, can be described by Tables 2.1 and 2.2.

Recall that for a valuation f , the maximum value of f is denoted by M_f , the set of points with f -value 0 is denoted by \mathcal{O}_f and the hyperplane formed by the set of points that have f -value less than M_f is denoted by H_f . The column "Value Distribution" in Table 2.1 contains arrays whose i -th element is the number of points of \mathcal{S} with f -value $i - 1$, for a valuation of the given row type. The columns of Table 2.2 contain the number of lines of the valuation geometry \mathcal{V} of a given type incident with an arbitrary point of the type of the column. For example, the second entry in the first column tells us that through each valuation of type A , there is exactly 1 line of \mathcal{V} that has type ABB .

Type	#	M_f	$ \mathcal{O}_f $	$ H_f $	Value Distribution
A	9	2	1	5	[1, 4, 4]
B	6	1	3	3	[3, 6, 0]

 Table 2.1.: The valuations of $\text{GQ}(2, 1)$

Type	A	B
AAA	2	—
ABB	1	3
BBB	—	1

 Table 2.2.: The lines of the valuation geometry of $\text{GQ}(2, 1)$

Now for any near polygon \mathcal{S}' which contains \mathcal{S} as a full subgeometry and has three points on each line, we can associate points and lines of \mathcal{S}' to the points and lines of \mathcal{V} by using Theorems 2.2.3, 2.2.6 and 2.3.2, unless all three valuations induced by points on a line of \mathcal{S}' are equal. We assign a type to the points of \mathcal{S}' based on the type of the points of \mathcal{V} they induce. The type of a line of \mathcal{S}' is simply a sorted string of the types of the points contained in the line. Thus, every point of \mathcal{S}' has type A or B , and every line of \mathcal{S}' has type AAA , ABB or BBB .

Lemma 2.3.3. *Let \mathcal{S}' be a generalized quadrangle which contains \mathcal{S} as a full subgeometry. Then*

- (1) *every point of \mathcal{S}' is at distance at most 1 from \mathcal{S} ;*
- (2) *every point of \mathcal{S}' which is contained in \mathcal{S} has type A ;*
- (3) *every point of \mathcal{S}' which is not contained in \mathcal{S} has type B and is collinear with precisely three points of \mathcal{S} forming a 1-ovoid.*

- Proof.* (1) Let x be a point of \mathcal{S}' . If $d(x, \mathcal{S}) \geq 2$, then by Lemma 2.2.4, $M_{f_x} \leq 0$, which is a contradiction.⁶
- (2) Since \mathcal{S} is a full subgeometry, by Lemma 1.2.5 it is isometrically embedded. Thus, for every point x of \mathcal{S} the valuation f_x is classical.
- (3) Let x be a point of \mathcal{S}' which is not contained in \mathcal{S} , and hence it is at distance 1 from \mathcal{S} . Then by Lemma 2.2.4, $M_{f_x} \leq 1$, and hence x must be of type B by Table 2.1. Moreover, the points of \mathcal{S} collinear with x are in bijection with the set \mathcal{O}_{f_x} , and therefore x is collinear with exactly three points of \mathcal{S} , which form a 1-ovoid. □

Corollary 2.3.4. *Let \mathcal{S}' be a generalized quadrangle containing \mathcal{S} as a full subgeometry. Then for every line $L = \{x_1, x_2, x_3\}$ of \mathcal{S}' , the valuations f_{x_1}, f_{x_2} and f_{x_3} of \mathcal{S} are all mutually distinct.*

Proof. If L is contained in \mathcal{S} , then these are three distinct classical valuations; so assume that L is not contained in \mathcal{S} . Then by Lemma 2.3.3, L either intersects \mathcal{S} in a unique type A point, or it lies at distance 1 from \mathcal{S} and has only type B points. In the former case, one of the three valuations is of type A while the rest are of type B , and hence all of them are distinct (note that $f_{x_1} = f_{x_2} * f_{x_3}$). In the latter case, we see from Theorem 2.3.2(b) that the valuations are distinct. □

By Corollary 2.3.4, every line $\{x_1, x_2, x_3\}$ of \mathcal{S}' gives rise to a line $\{f_{x_1}, f_{x_2}, f_{x_3}\}$ of the valuation geometry \mathcal{V} .

Lemma 2.3.5. *Let \mathcal{S}' be a generalized quadrangle of order $(2, 2)$ containing \mathcal{S} as a full subgeometry. Then the map $x \mapsto f_x$ between the points of \mathcal{S}' and the points of \mathcal{V} is a bijection.*

Proof. By Proposition 1.2.4, \mathcal{S}' has 15 points, 6 of which belong to \mathcal{S} . Since \mathcal{S} is isometrically embedded by Lemma 1.2.5, each of the 6 points of \mathcal{S} are in bijective correspondence with the 6 type A points of \mathcal{V} . The 9 points of \mathcal{S}' which are not in \mathcal{S} are at distance 1 from \mathcal{S} and each of them is collinear with exactly three points of \mathcal{S} forming a 1-ovoid by Lemma 2.3.3. Since there are exactly 9 points of type B in \mathcal{V} (see Table 2.1), it suffices to show that two distinct points of \mathcal{S}' which are not contained in \mathcal{S} induce distinct valuations. Let x, y be two such distinct points. Suppose $f_x = f_y = f$ and let z be a point of \mathcal{O}_f , which is then collinear to both x and y . Then z is incident with two lines in \mathcal{S} and since the order of \mathcal{S} is $(2, 2)$, it is incident with at most one other line. Thus, $\{x, y, z\}$ is a line of \mathcal{S}' . Now let z' be any other point of \mathcal{O}_f . Then z' is collinear with both x and y , which contradicts **NP2**. □

Lemma 2.3.6. *Let \mathcal{S}' be a generalized quadrangle of order $(2, 2)$ containing \mathcal{S} as a full subgeometry. Then the map $\{x_1, x_2, x_3\} \mapsto \{f_{x_1}, f_{x_2}, f_{x_3}\}$ is a bijective map between the lines of \mathcal{S}' and the lines of \mathcal{V} that have type AAA or ABB .*

Proof. By Corollary 2.3.4 and Lemma 2.3.5, two distinct lines of \mathcal{S}' induce distinct lines of \mathcal{V} . By a double count using Table 2.2, there are 6 lines of type AAA and 9 lines of type ABB in \mathcal{V} . The lines contained in \mathcal{S} are of type AAA , and since there are 6 of them, they correspond bijectively to the type AAA lines of \mathcal{V} . Let x be a point of type B in \mathcal{S}' . Then

⁶Of course a more elementary reasoning based on the GQ axiom also proves this, but we prove it this way since a similar argument will appear in other cases later.

x is collinear with three points of \mathcal{S} by Lemma 2.3.3, giving rise to three lines through x , all necessarily of type AAB as they contain a point of type A . Since the order of \mathcal{S} is $(2, 2)$, there are no other lines through x . Therefore, there are no lines of type ABB in \mathcal{S}' . Since there are 9 lines in \mathcal{S}' besides those contained in \mathcal{S} , these must bijectively correspond to the type ABB lines of \mathcal{V} . \square

Corollary 2.3.7. *Up to isomorphism, there is at most one generalized quadrangle of order $(2, 2)$ containing \mathcal{S} as a full subgeometry.*

Proof. Every such GQ, if it exists, must be isomorphic to the subgeometry of \mathcal{V} formed by taking all the lines of type AAA and ABB . \square

We can in fact show that every GQ of order $(2, 2)$ contains \mathcal{S} as a full subgeometry. Take two disjoint lines $L_1 = \{x_{11}, x_{12}, x_{13}\}, L_2 = \{x_{21}, x_{22}, x_{23}\}$ in such a GQ and assume that for each $i \in \{1, 2, 3\}$ the point x_{1i} is collinear with the point of x_{2i} . For each i , let x_{3i} denote the third point on the line joining x_{1i} and x_{2i} . Then from **GQ2** it follows that x_{31}, x_{32} and x_{33} are pairwise collinear and mutually distinct. Therefore, $L_3 = \{x_{31}, x_{32}, x_{33}\}$ is a line, and we get a (3×3) -grid formed by these lines. Corollary 2.3.7 now shows that up-to isomorphism, there is at most one GQ of order $(2, 2)$. Since we already know that there exists a generalized quadrangle of order $(2, 2)$, the $W(2)$ quadrangle (see Figures 1.1 and 1.2), it must be unique.

2.4. Some valuation geometries

In this section we record the data for valuation geometries of the split Cayley hexagon $H(2)$, its dual $H(2)^D$, and the Hall-Janko near octagon **HJ**, that will be used later. This data was obtained using the computer algorithms described in Chapter 3.

Type	#	M_f	$ \mathcal{O}_f $	$ H_f $	Value Distribution
A	63	3	1	31	$[1, 6, 24, 32]$
B	252	3	1	47	$[1, 14, 32, 16]$
C	252	2	1	23	$[1, 22, 40, 0]$
D	1008	2	5	31	$[5, 26, 32, 0]$

Table 2.3.: The valuations of $H(2)^D$

Type	A	B	C	D
AAA	3	—	-	-
ABB	2	1	-	-
ACC	2	-	1	-
ADD	24	-	-	3
BBB	-	4	-	-
BCC	-	1	2	-
BDD	-	4	-	2
CCC	-	-	8	-
CCD	-	-	40	5
CDD	-	-	4	2
DDD	-	-	-	10

Table 2.4.: The lines of the valuation geometry of $H(2)^D$

Type	#	M_f	$ \mathcal{O}_f $	$ H_f $	Value Distribution
A	63	3	1	31	[1, 6, 24, 32]
B_1	126	2	1	23	[1, 22, 40, 0]
B_2	252	2	3	27	[3, 24, 36, 0]
B_3	504	2	4	29	[4, 25, 34, 0]
B_4	72	2	7	35	[7, 28, 28, 0]
B_5	378	2	9	39	[9, 30, 24, 0]
C	36	1	21	21	[21, 42, 0, 0]

Table 2.5.: The valuations of $H(2)$

Type	A	B_1	B_2	B_3	B_4	B_5	C
AAA	3	-	-	-	-	-	-
AB_1B_1	1	1	-	-	-	-	-
AB_2B_2	6	-	3	-	-	-	-
AB_3B_3	16	-	-	4	-	-	-
AB_4B_4	4	-	-	-	7	-	-
AB_5B_5	3	-	-	-	-	1	-
$B_1B_1B_1$	-	3	-	-	-	-	-
$B_1B_1B_2$	-	16	4	-	-	-	-
$B_1B_1B_5$	-	6	-	-	-	1	-
$B_1B_2B_4$	-	4	2	-	7	-	-
$B_1B_3B_3$	-	12	-	6	-	-	-
B_1B_3C	-	12	-	3	-	-	42
$B_2B_2B_2$	-	-	12	-	-	-	-
$B_2B_2B_5$	-	-	6	-	-	2	-
$B_2B_3B_3$	-	-	10	10	-	-	-
B_2CC	-	-	1	-	-	-	14
$B_3B_3B_5$	-	-	-	3	-	2	-
B_4B_4C	-	-	-	-	1	-	1
$B_5B_5B_5$	-	-	-	-	-	1	-
B_5CC	-	-	-	-	-	1	21

Table 2.6.: The lines of the valuation geometry of $H(2)$

Type	#	M_f	$ \mathcal{O}_f $	Value Distribution
A	315	4	1	[1,10,80,160,64]
B	630	3	1	[1,10,112,192,0]
C	3150	3	1	[1,26,128,160,0]
D	1008	2	5	[5,110,200,0,0]
E	2016	2	25	[25,130,160,0,0]

Table 2.7.: The valuations of HJ

Type	A	B	C	D	E
AAA	5	–	–	–	–
ABB	1	1	–	–	–
ACC	5	–	1	–	–
BBB	–	5	–	–	–
BBC	–	10	1	–	–
CCC	–	–	9	–	–
CDD	–	–	4	25	–
DDD	–	–	–	6	–
DEE	–	–	–	1	1
EEE	–	–	–	–	6

Table 2.8.: The lines of the valuation geometry of \mathbf{HJ}

3. Computing Hyperplanes, Ovoids and Valuations

3.1. Introduction

In this chapter we describe the computational techniques we have used in our work to compute hyperplanes, ovoids and valuations of near polygons. We will construct all the Tables given in Section 2.4 by using computer models of these near polygons. We will also give a computer-aided proof of nonexistence of distance-2 ovoids (which are equivalent to 1-ovoids) in the dual split Cayley hexagon $H(4)^D$. Furthermore, these techniques will be applied in Chapter 4 to obtain some results on intersections of hyperplanes in split Cayley hexagons and their duals which will help us prove non-existence of certain generalized hexagons (see Theorem 4.3.11).

The study of distance- j ovoids in generalized polygons was initiated by Thas, who investigated the existence of distance-2 ovoids in generalized quadrangles and distance-3 ovoids in generalized hexagons [135] (these were simply called ovoids). In their general form, distance- j ovoids were introduced by Offer and Van Maldeghem in [113]. These are important objects in finite geometry, with connections to perfect codes in distance-regular graphs [44], cores of graphs [43] and several other topics. Our focus is on distance-2 ovoids in the dual split Cayley hexagons $H(q)^D$, where q is a prime power. While for the split Cayley hexagon $H(q)$, existence of these ovoids is known for $q \in \{2, 3, 4\}$ [68], and the ovoids are completely classified [69, 117]; for $H(q)^D$, only the nonexistence for $q = 2$ and the existence for $q = 3$ (noting that $H(3) \cong H(3)^D$) is known. Therefore, our result on nonexistence of distance-2 ovoids in $H(4)^D$ is the natural next step.

In [113], it was shown that a distance-3 ovoid in a generalized octagon of order (s, t) can only exist if $s = 2t$. This directly implies that there are no distance-3 ovoids in the Ree-Tits octagons $GO(q, q^2)$, or their duals $GO(q^2, q)$, where q is an odd power of 2, except for the case of $GO(4, 2)$. We will show that this last remaining case does not have any distance-3 ovoids either¹.

To be able to understand the computer code given in this, the reader needs to be familiar with the computer algebra system GAP [82] and the mathematical software SageMath [74]. But we have tried to explain the basic ideas without referring to any specific software, so that one can implement all the algorithms in whichever language or software one is comfortable with.

¹After we solved this last “open case”, we found out that Brouwer had already verified it using a different computation [33]. Brouwer’s result is mentioned as a remark in a *liber amicorum* in Dutch, which leaves out some details of the used techniques. So, we still see some value in including this result in the thesis.

3.2. Definitions and Basic Techniques

We first recall the definition of hyperplanes and 1-ovoids, since they are going to be central to this chapter.

Definition. A *hyperplane* of a point-line geometry is a proper subset of the set of points with the property that every line is either contained in it or intersects it in a unique point. If a hyperplane does not contain any line, then it is called a *1-ovoid*.

Definition. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near polygon and let $d(\cdot, \cdot)$ denote the distance function of its collinearity graph.

- (a) A *partial distance- j ovoid* of \mathcal{S} is a set \mathcal{O} of points such that for every two distinct points x and y in \mathcal{O} we have $d(x, y) \geq j$.
- (b) A *distance- j ovoid* of \mathcal{S} is a partial distance- j ovoid \mathcal{O} such that (1) for every point a of \mathcal{S} there exists a point x of \mathcal{O} such that $d(a, x) \leq j/2$; (2) for every line L of \mathcal{S} there exists a point $x \in \mathcal{O}$ such that $d(L, x) \leq (j-1)/2$.

Remark. The definition of distance- j ovoids given in [113] looks different as it employs the distance function in the incidence graph, which is a common convention in the literature on generalized polygons. But, it can easily be shown using Proposition 1.1.1 that these two definitions are equivalent.

The *exact cover* problem in a hypergraph (V, E) asks for the existence of a subset S of E such that for every vertex v there exists a unique edge e in S which contains v . The dual of this problem is the *exact hitting set* problem where we need to find a subset O of V such that for every edge e there is a unique vertex v in O which is contained in E . As noted before, 1-ovoids in point-line geometries, where the lines can be treated as subsets of points, are equivalent to exact hitting sets in the corresponding hypergraph. It is clear from the definitions that in a near polygon distance-2 ovoids are equivalent to 1-ovoids. We will now show that in general, distance- j ovoids can be linked to exact hitting sets in certain hypergraphs.

Lemma 3.2.1.² Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a near $2n$ -gon. For any $i \in \{0, \dots, n\}$ and an element $a \in \mathcal{P} \cup \mathcal{L}$, let $\Gamma_{\leq i}(a)$ denote the set of points at distance at most i from a in the collinearity graph of \mathcal{S} . Let \mathcal{O} be a set of points and $j \in \{2, \dots, n\}$. Then

- (1) for j even, \mathcal{O} is a distance- j ovoid if and only if for all $L \in \mathcal{L}$ we have

$$|\Gamma_{\leq (j-2)/2}(L) \cap \mathcal{O}| = 1;$$

- (2) for j odd, \mathcal{O} is a distance- j ovoid if and only if for all $x \in \mathcal{P}$ we have

$$|\Gamma_{\leq (j-1)/2}(x) \cap \mathcal{O}| = 1.$$

Proof. We only prove the first case, when j is even, and note that the second part has a similar proof. Say \mathcal{O} is a distance- j ovoid and let $L \in \mathcal{L}$. Then by the definition of distance- j ovoids there exists a point x in \mathcal{O} such that $d(x, L) \leq (j-1)/2$, but since j is even and distances are integral we have $d(x, L) \leq (j-2)/2$. Say there was another

²One side of this Lemma was proved in [57, Lemma 2].

point $y \neq x$ in \mathcal{O} with $d(y, L) \leq (j-2)/2$. Then $d(x, y) \leq d(x, L) + d(y, L) + 1 = j-1$ which is a contradiction. Now say \mathcal{O} is a set of points such that for every line M we have $|\Gamma_{\leq (j-2)/2}(M) \cap \mathcal{O}| = 1$. Let x, y be two distinct points in \mathcal{O} . If $d(x, y) \leq j-1$, then there exists a line L in the path joining x to y such $d(x, L) \leq (j-2)/2$ and $d(y, L) \leq (j-2)/2$, which is not possible. Now let x be an arbitrary point of \mathcal{S} . Let L be any line through x , and let y be the unique point in \mathcal{O} such that $d(L, y) \leq (j-2)/2$. Then $d(x, y) \leq 1 + d(L, y) = j/2$. Let L be an arbitrary line of \mathcal{S} ; then by the assumption on \mathcal{O} there exists a point in \mathcal{O} at distance at most $(j-2)/2 \leq (j-1)/2$ from L . Therefore, \mathcal{O} is a distance- j ovoid. \square

Lemma 3.2.1 makes it clear that the existence of a distance- j ovoid in a near $2n$ -gon \mathcal{S} is equivalent to the existence of an exact hitting set in a hypergraph derived from the collinearity graph of \mathcal{S} . For j even the edges of this hypergraph are the subsets $\Gamma_{\leq (j-2)/2}(L)$ of \mathcal{P} where L is a line, and for j odd the edges of this hypergraph are the subsets $\Gamma_{\leq (j-1)/2}(x)$ where x is a point.

Note that the exact cover problem is a well-known NP-hard problem, and hence we cannot expect a “really efficient” algorithm for computing distance- j ovoids in general. Still, we can do better than brute force search. One of the fastest known algorithms for computing exact covers is *Knuth’s Dancing Links* [98]. It can be implemented from scratch in any programming language, or one can use some existing implementation. We will be using the standard implementation in SageMath [74], which is due to Carlo Hamalainen³. The following function written in SageMath will give us an iterator for all exact hitting sets in a hypergraph, or equivalently all 1-ovoids in the corresponding point-line geometry.

```
def ovoids(P,L):
    """
    Find all exact hitting sets (1-ovoids) in a hypergraph (geometry).

    Args:
    P -- the vertices (points) of the hypergraph (geometry)
    L -- the edges (lines) of the hypergraph (geometry)
    Returns:
    an iterator for exact hitting sets (1-ovoids) in the hypergraph (geometry).
    """
    map = dict() # to construct the dual problem of exact covers
    for p in P:
        map[p] = []
    for i in range(len(L)):
        for p in L[i]:
            map[p].append(i)
    E = [map[p] for p in P]
    for match in DLXCPP(E):
        yield [P[i] for i in match]
```

To see how one can use this function, we determine all 1-ovoids of a (3×3) -grid (this corresponds to the type B valuation given in Table 2.1).

³<https://carlo-hamalainen.net/blog/2008/3/1/an-exact-cover-solver-for-sage>

```

P = [1,2,3,4,5,6,7,8,9]
L = [[1,2,3],[4,5,6],[7,8,9],[1,4,7],[2,5,8],[3,6,9]]
O = ovoids(P, L)
print list(O)
      [[1, 5, 9], [1, 6, 8], [2, 4, 9], [2, 6, 7], [3, 4, 8], [3, 5, 7]]

```

Finally we mention that distance- j ovoids in near polygons also give rise to valuations as follows. Let \mathcal{O} be a distance- j ovoid of a near polygon $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ for j even. Then the function $f : \mathcal{P} \rightarrow \mathbb{N}$ defined by $f(x) = d(x, \mathcal{O})$ is a valuation of \mathcal{S} , with $M_f = j/2$. We will not encounter distance- j ovoidal valuations in this thesis, and we refer the interested reader to [64] for more details.

3.3. Computer Models of Geometries

Since we want to compute the valuation geometries of some “small” near polygons, and compute certain substructures like distance- j ovoids and hyperplanes in these near polygons, we will need computer models that we can work with. In this section we discuss how we can obtain the computer models of some small split Cayley hexagons, their duals, the Ree-Tits octagon $\text{GO}(2, 4)$, and the Hall-Janko near octagon HJ .

To construct a computer model of the split Cayley hexagons (and their duals), one can use the definition mentioned in Section 1.2.1 which involves a particular quadric in $\text{PG}(6, q)$. This is what the GAP [82] package `FinIng` [15] does. However, for the small cases that we consider in this thesis, we prefer a different construction. Observing that the automorphism groups of these geometries act *primitively* and *distance transitively* (see for example [40]) on both points and lines of the geometry, we can use the permutation representations given in the database of primitive permutation groups of GAP (works only for $q \in \{2, 3, 4\}$) or the permutation representations given in the ATLAS of finite simple group [54]⁴ (works for $q \in \{2, 3, 4, 5\}$) to construct the points and lines of these geometries. This method also works for other geometries, like the generalized hexagons of order $(q, 1)$ obtained from $\text{PG}(2, q)$ (for small q), the generalized octagons of order $(q, 1)$ (for small *even* q), the Ree-Tits octagons and for the Hall-Janko near octagon.

Since for characteristic not equal to 3 we have two non-isomorphic permutation representations of the automorphism group of the split Cayley hexagons, one on the points of split Cayley hexagon and another one on the points of its dual (which is non-isomorphic in this case), we need a way to distinguish between these two. An easy combinatorial method of doing so was given by Ronan in [120], which we have used to find out which permutation representation corresponds to the dual split Cayley. The following function, written in GAP, constructs a near polygon with permutation group G acting distance transitively on its points, and transitively on the lines. As input, the function only requires the permutation group G acting on the set $\{1, \dots, v\}$ for some integer v and an array `partition_sizes` whose i -th entry denotes the number of points at distance $i - 1$ from a fixed point in the geometry⁵. We assume that all these entries are distinct, a condition which is satisfied by all our examples. Since we have a near polygon, we can

⁴see <http://brauer.maths.qmul.ac.uk/Atlas/v3/exc/>

⁵the index of an array in GAP start from 1

construct a line (and hence all lines via transitivity of the group) by taking the intersection $x^\perp \cap y^\perp$ of two collinear points x and y (p^\perp is equal to the set of points at distance at most 1 from p for any point p).

```

construct_geometry := function(G, partition_sizes)
  local v, i, j, k, temp, points, orbs, perp1, perp2, lines, partition, D, d;
  v := Sum(partition_sizes);
  points := [1..v];
  orbs := Orbits(Stabilizer(G,1),points);;
  partition := [];
  for i in partition_sizes do
    Add(partition, Filtered(orbs, x -> Size(x) = i)[1]);
    # this is where we use the assumption that the entries are distinct
  od;
  perp1 := Union([1],partition[2]);
  perp2 := OnSets(perp1,RepresentativeAction(G,1,partition[2][1]));
  lines := Orbit(G, Intersection(perp1, perp2), OnSets);
  D := NullMat(v, v);
  for i in [1..v] do
    for j in [i+1..v] do
      k := j^RepresentativeAction(G,i,1);
      temp := 0;
      while not(k in partition[temp+1]) do temp := temp + 1; od;
      D[i][j] := temp; D[j][i] := temp;
    od;
  od;
  d := function(x,y)
    return D[x][y];
  end;
  return [points, lines, d];
end;

```

Using this function, we can construct the computer models of all the geometries we will need in this thesis. We include the GAP code that can be used along with the function defined above to make these models. Recall that the number of points in $H(q)$ is $(1+q)(1+q^2+q^4)$ and its automorphism group has size $rq^6(q^6-1)(q^2-1)$ where $q = p^r$ for some prime p . The automorphism group of the Hall-Janko near octagon HJ is $J_2:2$ and has size 1209600, which acts primitively and distance-transitively on its 315 points with suborbit sizes (the entries of `partition_sizes`) 1, 10, 80, 160 and 64. The automorphism group of the Ree-Tits octagon $GO(2,4)$ has size 35942400, and it acts primitively and distance-transitively on its 1755 points with suborbit sizes 1, 10, 80, 640, 1024.

```

# H(2)
G := AllPrimitiveGroups(DegreeOperation, 63, Size, 12096)[2];
H2 := construct_geometry(G, [1, 6, 24, 32]);

# H(2)^D

```

```

G := AllPrimitiveGroups(DegreeOperation, 63, Size, 12096)[1];
H2D := construct_geometry(G, [1, 6, 24, 32]);

# HJ
G := AllPrimitiveGroups(DegreeOperation, 315, Size, 1209600)[1];
HJ := construct_geometry(G, [1, 10, 80, 160, 64]);

# H(3)
G := AllPrimitiveGroups(DegreeOperation, 364, Size, 4245696)[1];
H3 := construct_geometry(G, [1, 12, 108, 243]);

# H(4)
G := AllPrimitiveGroups(DegreeOperation, 1365, Size, 503193600)[1];
H4 := construct_geometry(G, [1, 20, 320, 1024]);

# H(4)^D
G := AllPrimitiveGroups(DegreeOperation, 1365, Size, 503193600)[2];
H4D := construct_geometry(G, [1, 20, 320, 1024]);

# G0(2, 4)
G := AllPrimitiveGroups(DegreeOperation, 1755, Size, 35942400)[1];
Ree-Tits := construct_geometry(G, [1, 10, 80, 640, 1024]);
    
```

Finally we note that there is a database of small generalized polygons maintained by Moorhouse [110] that one can use to construct these computer models.

3.4. Valuation Geometries

While ovoidal, semi-classical and distance- j ovoidal valuations of a near polygon \mathcal{S} can in principle be computed using the Dancing Links algorithm in a reasonably fast way, we do not have any efficient way of doing so for other valuations (except of course the classical valuations). The case when every line of \mathcal{S} has three points is different. Then we have the following characterization of hyperplanes, which can be turned into an algorithm for computing valuations and valuation geometries of these near polygons.

Proposition 3.4.1. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a (finite) partial linear space with three points on each line and let $v = |\mathcal{P}|$. For any subset S of \mathcal{P} , let $\chi(S)$ denote the binary vector in the vector space \mathbb{F}_2^v whose i -th entry is 1 if the i -th point (after fixing an ordering of \mathcal{P}) of \mathcal{S} lies in S and 0 otherwise. Then a proper subset H of \mathcal{P} is a hyperplane of \mathcal{S} if and only if $\chi(\mathcal{P} \setminus H) \neq 0$ is orthogonal to $\chi(L)$ for every $L \in \mathcal{L}$ with respect to the standard inner product on \mathbb{F}_2^v .*

Proof. Let L be a line of \mathcal{S} . Since $|L| = 3$, the statement “ L is either contained in H or intersects H in exactly one point” is equivalent to $|L \cap H| \in \{1, 3\}$. Therefore, H is a hyperplane if and only if for all lines L , we have $|L \cap (\mathcal{P} \setminus H)| \in \{0, 2\}$ which is equivalent to $\chi(L) \cdot \chi(\mathcal{P} \setminus H) = 0$ in \mathbb{F}_2^v . \square

Corollary 3.4.2. *Let \mathcal{H} be the set of hyperplanes of a (finite) near polygon $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ that has three points on each line. Then the set $\{\chi(\mathcal{P} \setminus H) \mid H \in \mathcal{H}\} \cup \{0\}$ is equal to the null space of the $|\mathcal{L}| \times |\mathcal{P}|$ incidence matrix of \mathcal{S} over \mathbb{F}_2 .*

By Corollary 3.4.2, computing the set of hyperplanes of near polygons with three points on each line is equivalent to a computation of the null space of a binary matrix.⁶ Instead of storing all the elements of this null space, we would like to store only the non-isomorphic ones under the action of the automorphism group of the near polygon. To do so we use Linton's smallest image set algorithm [105], which has been implemented in the GRAPE package [130] of GAP. The function `SmallestImageSet(H, S)` returns the lexicographically smallest element of the orbit of the set S under the action of the group H . For the rest of this section assume that we are only talking about near polygons with three points on each line.

In the following GAP code, we compute the distinct representatives of the isomorphism classes of hyperplane complements of a near polygon \mathcal{S} that has three points on each line. It requires the automorphism group, points, lines, diameter and the distance function of \mathcal{S} , stored in the variables `G`, `points`, `lines`, `diam` and `dist`, respectively. We also assume that G acts transitively on the points so that all singular hyperplanes are isomorphic. To speed up the process we start with the singular hyperplanes, for which we have a variable `singular_compl` that contains the set of points opposite to a fixed point.

```

SetVect := function(S)
#I: subset of points
#O: characteristic vector of S
return One(GF(2))*List(points,function(i) if i in S then return 1;
    else return 0; fi; end);
end;

VectSet:=function(v)
#I: characteristic vector
#O: subset of points corresponding to v
return Set(Filtered(points,i->v[i]=One(GF(2))));
end;

IncidenceMatrix := TransposedMat(List(lines, SetVect));;
Y := NullspaceMat(IncidenceMatrix);;
U := Subspace(GF(2)^(Size(points)),Y);;

LoadPackage("grape");
hypcomplements := NewDictionary(Set([1,2,3]), false);
singular_compl := Filtered(points, x -> dist(x, 1) = diam);
AddDictionary(hypcomplements, SmallestImageSet(G,Set(singular_compl)));
balance := 2^(Dimension(U))- 1 - Size(points);

while balance > 0 do
w := Random(U);

```

⁶I learned this trick from Bart De Bruyn, who in turn learned it from Sergey Shpectorov.

```

if w <> Zero(U) then
  compl := VectSet(w);
  S := SmallestImageSet(G, compl);
  if not KnowsDictionary(hypcomplements, S) then
    AddDictionary(hypcomplements, S);
    balance:=balance-Index(G,Stabilizer(G,S,OnSets));
  fi;
fi;
od;

hypcomplements := hypcomplements!.list;

```

At the end of this computation we will have a list of all hyperplane complements, and hence hyperplanes of the near polygon \mathcal{S} . As an example, let \mathcal{S} be isomorphic to $H(2)$. Then computation done by the code above quickly shows that \mathcal{S} has 25 distinct hyperplanes up to isomorphism. This number coincides with that obtained by Frohardt and Johnson [81] who classified all hyperplanes of $H(2)$ and $H(2)^D$.

From these hyperplanes, we would now like to construct the valuation geometry of \mathcal{S} . Firstly, while every valuation f of \mathcal{S} gives rise to a hyperplane H_f , not every hyperplane is going to give rise to a valuation of \mathcal{S} . An example of this situation is given in Figure 1.2 where the subgrid is a hyperplane which is not of valuation type (try to see this using the Algorithm described below). So, we will pick a hyperplane complement and use it to try to construct a semi-valuation in which every point in the hyperplane complement has a constant value and every point not in the hyperplane complement has value strictly less than that constant. Note that once we have assigned values to two points of a line of \mathcal{S} , then the value of the third point is automatically determined. Using these ideas, we have Algorithm 1. We note that this idea of constructing valuations is due to De Bruyn [62].

We define a *partial valuation* of \mathcal{S} as a function f defined on a subspace S_f of \mathcal{S} such that f is a semi-valuation on the subgeometry of \mathcal{S} induced on S_f . In Algorithm 1 we use a function *AssignValue* which takes as input a function f defined on a set S of points, a point x and a value i , and finds a partial valuation g such that: (i) S_g is the smallest subspace containing S and $\{x\}$; (ii) $g(y) = f(y)$ for all $y \in S$; (iii) $g(x) = i$. If such a function g does not exist then *AssignValue* returns *fail*. An implementation of this function and Algorithm 1 in GAP will follow. Note that we start with assigning value 0 to all points of a given hyperplane complement, and thus every other point will get a value from the set $\{-1, -2, \dots, -n\}$ where n is the diameter of \mathcal{S} .

We now give the GAP code based on Algorithm 1 which we have used to compute all valuations.

```

MakeClosure := function(val)
  local X,x,L,f,S,V;
  f := ShallowCopy(val);
  X := Filtered(points, x -> val[x] <> fail);
  for L in lines do
    S := Intersection(L,X);
    if Size(S) = 2 then

```

Algorithm 1 Create valuations

input: a hyperplane complement H *output:* list of all semi-valuations f for which $\overline{H_f} = H$ and for which the maximal value is equal to 0 $complete \leftarrow []$ $incomplete \leftarrow []$ $val \leftarrow$ function which is 0 on all points of H and undefined on other points $f = AssignValue(val, H[1], 0)$ **if** f is *fail* **then** **return** *fail***end if****if** f is defined on all points **then** Add f to *complete***else** Add f to *incomplete***end if****while** *incomplete* is not empty **do** $val \leftarrow$ an element popped from *incomplete* $x \leftarrow$ a random point for which val is not defined **for** i in $\{-1, -2, \dots, -n\}$ **do** $f \leftarrow AssignValue(val, x, i)$ **if** f is *fail* **then** **return** *fail* **end if** **if** f is defined on all points **then** Add f to *complete* **else** Add f to *incomplete* **end if** **end for****end while****return** *complete*

```
x := Difference(L,X)[1];
if AbsInt(val[S[1]] - val[S[2]]) > 1 then return [false,val]; fi;
if val[S[1]] = val[S[2]] then f[x] := val[S[1]] - 1;
else f[x] := Maximum(val[S[1]], val[S[2]]); fi;
fi;
if Size(S) = 3 then
  V := val{S};
  Sort(V);
  if not ((V[1] = V[2] - 1) and (V[2] = V[Size(V)]))
    then return [false, val];
  fi;
fi;
od;
return [true,f];
end;
```

```
AssignValue := function(val, x, i)
  local f,g,temp;
  f := [];
  g := ShallowCopy(val);
  g[x] := i;
  while f <> g do
    f := ShallowCopy(g);
    temp := MakeClosure(g);
    if not temp[1] then return [false, val]; fi;
    g := temp[2];
  od;
  return [true,f];
end;
```

```
AreIsomorphicVal := function(val1,val2)
  return val2 in Set(G, g -> List(points, x -> val1[x^g]));
end;
```

```
CreateVal := function(compl)
#I: a hyperplane complement compl
#O: all possible valuations corresponding to compl
  local i, p, x, q, X, temp, complete, incomplete,val;
  complete := [];
  incomplete := [];
  val := ListWithIdenticalEntries(Size(points), fail);
  for x in compl do val[x] := 0; od;
  temp := AssignValue(val, compl[1], 0);
  if not temp[1] then return []; fi;
  if ForAll(points, x -> temp[2][x] <> fail) then Add(complete, temp[2]);
  else Add(incomplete, temp[2]);
  fi;
fi;
```

```

while incomplete <> [] do
  val := Remove(incomplete);
  p := Filtered(points, x -> val[x] = fail)[1];
  for i in [1..diam] do
    temp := AssignValue(val, p, -i);
    if temp[1] then
      if ForAll(points, x -> temp[2][x] <> fail) then
        if not ForAny(complete, f -> AreIsomorphicVal(f, temp[2])) then
          Add(complete, temp[2]);
        fi;
      else Add(incomplete, temp[2]);
      fi;
    fi;
  od;
od;
return Set(complete, L -> List(L, x -> x - Minimum(L)));
end;

Valuations := [];
for h in hypcomplements do
  S := CreateVal(h);
  if S <> [] then
    Append(Valuations, S);
  fi;
od;

```

Once we have all the valuations of \mathcal{S} , we can easily construct the first table in the description of the valuation geometry \mathcal{V} of \mathcal{S} . For the second table, and to verify other properties of \mathcal{V} , we still need the lines of \mathcal{V} . To construct the lines, we simply go through pairs $\{f_1, f_2\}$ of all valuations, check if they are neighboring, and if they are then compute $f_1 * f_2$ to construct the line $\{f_1, f_2, f_1 * f_2\}$ of \mathcal{V} . We do so in the following GAP code. We assume that there is a user defined function **TypeVal** which takes as input a valuation and returns its type, which can be any string that identifies the different non-isomorphic valuations. For example, we use the following function for $H(2)^D$ (see Table 2.3).

```

TypeVal := function(val)
  local h, M, x;
  M := Maximum(val);
  h := Filtered(points, x -> val[x] = M);
  if Size(h) = 32 and M = 3 then return 'A'; fi;
  if Size(h) = 40 then return 'C'; fi;
  if Size(h) = 16 then return 'B'; fi;
  return 'D';
end;

```

As noted in Section 2.4, these point types are used to give types to the lines of \mathcal{V} . In the following code, we give some functions that can be used to find the types of all lines

through a fixed point of \mathcal{V} , which is what we need to construct the second table that describes the valuation geometry \mathcal{V} .

```
Isneighboring := function(val1, val2)
#I: two valuations val1 and val2
#0: returns whether they are distinct and neighboring
  local S;
  if val1 = val2 then return false; fi;
  S := List(points, x -> val1[x] - val2[x]);
  Sort(S);
  if S[Size(S)] - S[1] > 2 then return false; fi;
  return true;
end;

Epsilon := function(val1, val2)
#I: two distinct neighboring valuations val1 and val2
#0: the integer e for which  $|\text{val1}(x) - \text{val2}(x) + e| \leq 1$ 
  local e, p, L, x1, x2;
  for L in lines do
    x1 := Filtered(L, x -> val1[x] = Minimum(val1{L}))[1];
    x2 := Filtered(L, x -> val2[x] = Minimum(val2{L}))[1];
    if x1 <> x2 then break; fi;
  od;
  return val2[x2] - val1[x1];
end;

ThirdFromTwo := function(val1, val2)
#I: two valuations val1 and val2
#0: the new valuation val1*val2
  local e, x, val3;
  e := Epsilon(val1, val2);
  val1 := val1 + e;
  val3 := 0*points;
  for x in points do
    if val1[x] = val2[x] then val3[x] := val1[x] - 1;
    else val3[x] := Maximum([val1[x], val2[x]]);
  fi;
od;
  return val3 - Minimum(val3);
end;

ThroughPoint := function(s, S)
#I: a member s of the list Valuations and a list S of some valuations
#0: the types of lines through point s of the valuation geometry
#   when the other two points (valuations) lie in set S
  local val1, val2, val3, L, types, x, g, i, T;
  val1 := s;
  types := Set([]);
```

```

for val2 in S do
  if Isneighboring(val1, val2) then
    val3 := ThirdFromTwo(val1, val2);
    if val3 in S then
      L := [val1, val2, val3];
      T := [TypeVal(val1), TypeVal(val2), TypeVal(val3)]; Sort(T);
      i := Position(List(types, x -> x[1]), T);
      if i <> fail then types[i][2] := types[i][2] + 1/2;
      else AddSet(types, [T, 1/2]); fi;
    fi;
  fi;
od;
return types;
end;

```

```

GenerateOrbit := function(s)
#I: a valuation s
#O: the orbit of this valuation under the action of group G
  return Set(G, g -> List(points, x -> s[x^Inverse(g)]));
end;

```

```

Table2 := [];
S := Set([]);
for s in Valuations do S := Union(S, GenerateOrbit(s)); od;
for s in Valuations do
  Add(Table2, [TypeVal(s), ThroughPoint(s, S)]);
od;

```

After this computation, the variable `Table2` contains a list whose i -th entry corresponds to a type of the valuation and the number of lines of each type through a valuation of that type. For example, when \mathcal{S} is isomorphic to $H(2)^D$ we have the following entries (see Table 2.4):

```

['C', [['ACC', 1], ['BCC', 2], ['CCC', 8], ['CCD', 40], ['CDD', 4]]]
['D', [['ADD', 3], ['BDD', 2], ['CCD', 5], ['CDD', 2], ['DDD', 10]]]
['A', [['AAA', 3], ['ABB', 2], ['ACC', 2], ['ADD', 24]]]
['B', [['ABB', 1], ['BBB', 4], ['BCC', 1], ['BDD', 4]]]

```

3.5. Distance-2 ovoids in $H(4)^D$

While it can be directly shown using the Dancing Links algorithm that $H(2)^D$ does not have any distance-2 ovoids, for $H(4)^D$ this does not work. We can use the automorphism group of $H(4)^D$ to perform some sort of symmetry breaking (see for example [117]), but even then it does not seem likely that we will get any results in a reasonable time frame. The main idea that makes this computational problem feasible is as follows:

First classify all distance-2 ovoids in a fixed subhexagon isomorphic to $H(4, 1)$ up to isomorphism under the action of the stabilizer of $H(4, 1)$, and then see if any of these ovoids can be extended to a distance-2 ovoid of $H(4)^D$.

We have implicitly assumed that every distance-2 ovoid of $H(4)^D$ will induce a distance-2 ovoid of its subhexagon isomorphic to $H(4, 1)$. This is justified by the following lemma.

Lemma 3.5.1. *Let \mathcal{S} be a generalized hexagon of order (s, t) . Let \mathcal{H} be a subhexagon of \mathcal{S} that has order (s, t') and let \mathcal{O} be a distance-2 ovoid of \mathcal{S} . Then $\mathcal{O} \cap \mathcal{H}$ is a distance-2 ovoid of \mathcal{H} and*

$$|\mathcal{O} \cap \mathcal{H}| = s^2 t'^2 + s t' + 1.$$

Proof. By Lemma 3.2.1, \mathcal{O} is a distance-2 ovoid if and only if it meets every line in a unique point. If each line of \mathcal{S} meets \mathcal{O} in exactly 1 points, then the same is true for \mathcal{H} . Moreover, the number of lines in a generalized hexagon of order (s, t) is $(1+t)(1+st+s^2t^2)$, and thus by double counting, the number of points in a distance-2 ovoid is $(1+st+s^2t^2)$. Therefore, we have $|\mathcal{O} \cap \mathcal{H}| = s^2 t'^2 + s t' + 1$. \square

We note that the stabilizer of a subgeometry of $H(4)^D$ which is isomorphic to $H(4, 1)$, under the action of the automorphism group of $\text{Aut}(H(4)^D)$ is in fact isomorphic to $\text{Aut}(H(4, 1))$. From the definition of $H(q, 1)$ as the dual of the incidence graph of $\text{PG}(2, q)$, it directly follows that a distance-2 ovoid in $H(q, 1)$ corresponds to a perfect matching of the incidence graph of $\text{PG}(2, q)$. It is folklore that the number of perfect matchings in a balanced bipartite graph corresponds to the permanent of the biadjacency matrix of that graph. The following can be verified by computing the corresponding permanent.

Lemma 3.5.2 ([127, A000794]). *The number of perfect matchings in the incidence graph of $\text{PG}(2, 4)$ is 18534400.*

Since a perfect matching of a bipartite graph is nothing but an exact cover, we can use Knuth's Dancing Links algorithm to enumerate all of these 18534400 distance-2 ovoids of $H(4, 1)$, and then by using Linton's smallest image set algorithm in exactly the same way as done in Section 3.4 we can find out the number of distinct distance-2 ovoids up to isomorphism.

Proposition 3.5.3. *Let G be the automorphism group of $H(4)^D$. Let \mathcal{H} be a subhexagon of $H(4)^D$ isomorphic to $H(4, 1)$. Then there are exactly 350 non-isomorphic distance-2 ovoids in \mathcal{H} with respect to $G_{\mathcal{H}}$, the stabilizer of \mathcal{H} under the action of G .*

Once we have computed these 350 partial distance-2 ovoids of $H(4, 1)$, it remains to show that none of them can be extended to a full distance-2 ovoid. For this, we use the following approach suggested by F. Ihringer which involves Integer Linear Programming (ILP). Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \text{I})$ be a near $2n$ -gon. Let \mathcal{O}' be a possibly empty set of points which forms a partial distance- j ovoid for some $j \in \{2, \dots, n\}$, i.e., every pair of points in \mathcal{O}' are at distance at least j in the point graph. Let $H = (V, E)$ be the hypergraph as defined above, with $V = \mathcal{P}$ and

$$E = \begin{cases} \{\Gamma_{\leq (j-2)/2}(L) : L \in \mathcal{L}\} & \text{if } j \text{ is even} \\ \{\Gamma_{\leq (j-1)/2}(p) : p \in \mathcal{P}\} & \text{if } j \text{ is odd.} \end{cases}$$

For each $p \in \mathcal{P}$ let $X_p \in \{0, 1\}$ be a binary variable. Then the equations

$$\begin{aligned} X_p &= 1 && \text{for all } p \in \mathcal{O}' \\ \sum_{p \in e} X_p &= 1 && \text{for all } e \in E \end{aligned} \quad (3.1)$$

have an integer solution if and only if \mathcal{S} possesses a distance- j ovoid that contains \mathcal{O}' . Similarly, the equations

$$\begin{aligned} X_p &= 1 && \text{for all } p \in \mathcal{O}' \\ \sum_{p \in e} X_p &\leq 1 && \text{for all } e \in E \end{aligned} \quad (3.2)$$

have an integer solution if and only if \mathcal{S} possesses a partial distance- j ovoid that contains \mathcal{O}' .

Once we have cast the given problem as such an integer linear programming (ILP), we can use fast ILP solvers like Gurobi and CPLEX to solve the original problem. In fact, if we take $\mathcal{O}' = \emptyset$, then this method can alternately be used to test existence of distance- j ovoids in near polygons. This is what we do for the case of $\text{GO}(4, 2)$, where by using Gurobi we have verified that there is no distance-3 ovoids.⁷

For each of the 350 partial distance-2 ovoids \mathcal{O}' of $\text{H}(4)^D$ we can now define a integer linear optimization problem as in (3.1). Then the ILP solvers quickly show that these equations are infeasible for all of the 350 cases.⁸ Our full code for this computation is available online⁹.

Remark. For the next open case, $\text{H}(5)^D$, our algorithmic approach fails for several reasons:

- (1) The incidence graph of $\text{PG}(2, 5)$ has 4598378639550 perfect matchings while the automorphism group of $\text{PG}(2, 5)$ has size only 744000. So a classification of all non-isomorphic distance-2 ovoids of $\text{H}(5, 1)$ seems out of reach via current methods.
- (2) Even for one given distance-2 ovoid of $\text{H}(5, 1)$, the corresponding integer linear program takes too long to solve with state-of-the-art ILP solvers.

Finally we note that our techniques can also be used to give upper bounds on the size of a partial distance-2 ovoid in $\text{H}(q)^D$, via the following lemma.

Lemma 3.5.4. *Let \mathcal{O} be a partial distance-2 ovoid of $\text{H}(q)^D$. Suppose that no subhexagon \mathcal{H} of $\text{H}(q)^D$ isomorphic to $\text{H}(q, 1)$ contains $q^2 + q + 1$ points of \mathcal{O} , which necessarily form a distance-2 ovoid of \mathcal{H} . Then $|\mathcal{O}| \leq (q^2 - q + 1)(q^2 + q)$*

Proof. We double count (p, \mathcal{H}) , where \mathcal{H} a subhexagon of $\text{H}(q)^D$ isomorphic to $\text{H}(q, 1)$ and $p \in \mathcal{O} \cap \mathcal{H}$. From [138, Corollary 1.8.6] it follows that through every pair of opposite lines in $\text{H}(q)^D$, there is a unique subhexagon isomorphic to $\text{H}(q, 1)$. Via an easy counting

⁷The running time was about one day with Gurobi Optimizer version 6.5.0 build v6.5.0rc1 (linux64) with an Intel Core i5-3550 CPU @ 3.30GHz processor.

⁸We verified this with CPLEX (several versions), Gurobi Optimizer (several versions) and the constraint solver Minion. The 350 ILPs in 350 files in the LP format took 540.3 seconds with Gurobi Optimizer version 6.5.0 build v6.5.0rc1 (linux64) with an Intel Core i5-3550 CPU @ 3.30GHz processor. Minion's running times were similar.

⁹<http://math.ihringer.org/data.php>

argument, we then know that each point is contained in $(1+q)q^3/2$ subhexagons isomorphic to $H(q, 1)$, which tells us that there are $|\mathcal{O}|(1+q)q^3/2$ such pairs. We also know that there are $q^3(1+q)(q^2-q+1)/2$ subhexagons of $H(q)^D$ which are isomorphic to $H(q, 1)$. Under the condition $|\mathcal{O} \cap \mathcal{H}| \leq q^2 + q$ this yields $|\mathcal{O}| \leq (q^2 - q + 1) \cdot (q^2 + q)$. \square

For $q = 2$, Lemma 3.5.4 gives us $|\mathcal{O}| \leq 18$ and for $q = 4$ it gives us $|\mathcal{O}| \leq 260$ under the given assumption that \mathcal{O} does not intersect a given subhexagon in $q^2 + q + 1$ points. We can now use the following computational approach. If the ILP defined in (3.2) does not have a solution larger than some integer $b \geq (q^2 - q + 1)(q^2 + q)$ for all non-isomorphic distance-2 ovoids of $H(q, 1)$, then we obtain b as an upper bound on the size of a partial distance-2 ovoids. Using this approach, so far we have been able to show that every partial distance-2 ovoid in $H(4)^D$ has size at most 265. With more time, we might be able to determine the exact bound in this case.

4. Generalized Hexagons Containing a Subhexagon

4.1. Introduction

Tits posed the following question about generalized polygons (see [138, Section 1.7.8]):

Does there exist a generalized $2n$ -gon, $n \geq 2$, of order (s, ∞) where $s \in \mathbb{N} \setminus \{0, 1\}$ and ∞ is any infinite cardinal?

Such a generalized polygon is called *semi-finite*. Note that by Proposition 1.2.1(1), no such generalized $(2n+1)$ -gon can exist. This problem turned out to be extremely hard and the only progress that has been made so far is in the case of $n = 2$ and $s \in \{2, 3, 4\}$, where non-existence is proved by Cameron [41] for $s = 2$, Brouwer [32] and Kantor (unpublished) for $s = 3$, and Cherlin [46] for $s = 4$. A modified, and easier, version of the question that can be asked is as follows:

Given a finite generalized $2n$ -gon \mathcal{S} , $n \geq 2$, of order (s, t) , with $s \in \mathbb{N} \setminus \{0, 1\}$, does there exist a generalized $2n$ -gon of order (s, ∞) containing \mathcal{S} as a subgeometry?

In this chapter, we will solve this problem for $n = 3$ and $s, t \in \{2, 3, 4\}$ with $s = t$, when \mathcal{S} is one of the known generalized hexagons of order s , by proving non existence of a semi-finite generalized hexagon containing \mathcal{S} . In fact, we will prove much more in some specific cases. We will show that for $s \in \{2, 4\}$, if \mathcal{S} is isomorphic to the split Cayley hexagon $H(s)$, then there is no generalized hexagon with $s + 1$ points on each line which contains $H(s)$ as a proper subgeometry. And in Section 4.2 we will prove our results in the more general setting of near hexagons.

4.2. Near Hexagons of Order $(2, t)$

In this section we prove the following theorem by using valuation theory and the classification result of Cohen and Tits [53] which states that every finite generalized hexagon with 3 points on each line is isomorphic to $H(2, 1)$, $H(2)$, $H(2)^D$ or $T(2, 8)$.

Theorem 4.2.1. *Let \mathcal{S}' be a near hexagon with three points on each line containing a generalized hexagon \mathcal{S} of order 2 as an isometrically embedded subgeometry. Then the following holds:*

- (1) \mathcal{S}' is finite;

- (2) if \mathcal{S} is isomorphic to $H(2)^D$, then \mathcal{S} is a generalized hexagon and hence isomorphic to $H(2)^D$ or $T(2, 8)$;
- (3) if \mathcal{S} is isomorphic to $H(2)$ and \mathcal{S} is a generalized hexagon, then $\mathcal{S}' = \mathcal{S}$.

4.2.1. Near Hexagons Containing $H(2)^D$

For this section, let \mathcal{S}' be a near hexagon which contains $\mathcal{S} \cong H(2)^D$ isometrically embedded in it as a full and *proper* subgeometry. Let \mathcal{V} be the valuation geometry of \mathcal{S} . From Table 2.3 and the basic theory of valuations discussed in Chapter 2 we see that every point x of \mathcal{S}' has type A , B , C or D , depending on the type of the valuation f_x of \mathcal{S} it induces. If $L = \{x_1, x_2, x_3\}$ is a line of \mathcal{S}' , then $\{f_{x_1}, f_{x_2}, f_{x_3}\}$ is either a line of \mathcal{V} and thus has one of the 11 types mentioned in Table 2.4, or $f_{x_1} = f_{x_2} = f_{x_3}$ in which case L has the type XXX for some $X \in \{A, B, C, D\}$. We will sometimes refer to the points and lines of \mathcal{V} as \mathcal{V} -points and \mathcal{V} -lines.

Lemma 4.2.2. (1) *Every point of \mathcal{S}' has distance at most 1 from \mathcal{S} .*

(2) *Every point x of \mathcal{S} has type A , and the valuation f_x is classical with center x .*

(3) *Every point x at distance 1 from \mathcal{S} has type C and is collinear with a unique point x' of \mathcal{S} . Moreover, $\mathcal{O}_{f_x} = \{x'\}$.*

Proof. Let x be a point of \mathcal{S}' . If $d(x, \mathcal{S}) \geq 2$, then by Lemma 2.2.4, $M_{f_x} \leq 1$, which is not possible since every valuation of \mathcal{S} has maximum value at least 2 by Table 2.3. Therefore, every point is at distance at most 1 from \mathcal{S} . We have already seen in Theorem 2.2.3, that points of \mathcal{S} induce classical valuations of \mathcal{S} , and hence they are of type A .

Let x be a point not contained in \mathcal{S} . Then since $d(x, \mathcal{S}) = 1$, by Lemma 2.2.4 we have $M_{f_x} \leq 2$, and therefore x is either of type C or D (type A and B valuations have maximum value 3). Suppose that x is of type D . From Table 2.3 we see that there are five points of \mathcal{S} with f_x -value 0 giving rise to five points in \mathcal{S} collinear with x , and thus five lines through x intersecting \mathcal{S} in a point. By Table 2.4, each of these lines must have type ADD . The five lines of \mathcal{V} induced by these five lines of \mathcal{S}' are mutually distinct since the five classical valuations contained in them are mutually distinct (as their centers are distinct). A contradiction follows from the fact that through a given \mathcal{V} -point of type D , there are only three distinct \mathcal{V} -lines of type ADD (see the entry in row ADD and column D in Table 2.4). Therefore, x is of type C and thus $|\mathcal{O}_{f_x}| = 1$, which implies that x is collinear with a unique point x' of \mathcal{S} such that $\mathcal{O}_{f_x} = \{x'\}$. \square

Corollary 4.2.3. *Every point of \mathcal{S}' has type A or C and every line of \mathcal{S}' has type AAA , ACC or CCC . Moreover, every point of type C is incident with a unique line of type ACC .*

Corollary 4.2.4. *For every line $L = \{x_1, x_2, x_3\}$ of \mathcal{S}' the valuations f_{x_1} , f_{x_2} and f_{x_3} are mutually distinct and hence they form a line of \mathcal{V} .*

Proof. The proof is similar to that of Corollary 2.3.4. \square

Let ψ be the map between points and lines of \mathcal{S}' and points and lines of \mathcal{V} defined by $\psi(x) = f_x$ and $\psi(\{x_1, x_2, x_3\}) = \{f_{x_1}, f_{x_2}, f_{x_3}\}$. Then Corollary 4.2.3 implies that ψ maps a point of \mathcal{S}' to a \mathcal{V} -point of type A or C and maps a line of \mathcal{S}' to a \mathcal{V} -line of type AAA ,

ACC or *CCC*. Also note that every point of \mathcal{S}' which is not contained in \mathcal{S} , and hence lies in $\Gamma_1(\mathcal{S})$ by Lemma 4.2.2, is incident with a unique line that intersects \mathcal{S} and this line has type *ACC*.

Lemma 4.2.5. *Let \mathcal{V}' be the subgeometry of \mathcal{V} formed by taking the points of type *C* and the lines of type *CCC*. Then \mathcal{V}' is connected.*

Proof. This can be easily checked in the computer model of \mathcal{V} constructed via the code given in Chapter 2, as we just need to determine if a 16-regular graph on 252 vertices is connected (see Tables 2.3 and 2.4). \square

Lemma 4.2.6. *Suppose there are no quads in \mathcal{S}' that contain a point of \mathcal{S} . Then the following holds:*

- (1) *every point of $\Gamma_1(\mathcal{S})$ in \mathcal{S}' is incident with precisely nine lines, exactly one of which intersects \mathcal{S} ;*
- (2) *if x is a point of $\Gamma_1(\mathcal{S})$ and L_1, \dots, L_8 are the eight lines of type *CCC* through x , then the \mathcal{V} -lines $\psi(L_1), \dots, \psi(L_8)$ in \mathcal{V} are precisely the eight \mathcal{V} -lines of type *CCC* through f_x (see Table 2.4).*
- (3) *every valuation of type *C* is induced equally many times by the points of $\Gamma_1(\mathcal{S})$.*

Proof. (1) Let x be a point of $\Gamma_1(\mathcal{S})$ and x' the unique point of \mathcal{S} which is collinear with x . Let Y denote the set of neighbors of x which do not lie on the line xx' . For any $y \in Y$, let y' denote the unique point of \mathcal{S} collinear with y . The point y' , for any given $y \in Y$, must be noncollinear with x' as otherwise x' and y will have two common neighbors giving rise to a quad in \mathcal{S}' which intersects \mathcal{S} . Let Y' denote the set of all points of $\Gamma_2(x) \cap \mathcal{S}$ which are non-collinear with x' . Then we have shown that $y' \in Y'$ for all $y \in Y$.

The points of $\Gamma_2(x) \cap \mathcal{S}$ are precisely those points of \mathcal{S} that have f_x -value 1, and by Table 2.3, there are precisely 22 of them. Out of these 22, 6 are neighbors of x' and hence $|Y'| = 16$. We now show that the map $\varphi : Y \mapsto Y'$ where $\varphi(y) = y'$ is a bijection.

For every $z \in Y'$, there exists a unique common neighbor y of x and z , since if there are more than one, then we will get a quad intersecting \mathcal{S} . This y cannot lie in \mathcal{S} since otherwise it will coincide with the unique neighbor of x in \mathcal{S} , the point x' , which will contradict the fact that the points of Y' are noncollinear with x' . Therefore, φ is a bijection. This proves that $|Y| = 16$ and thus the total number of lines of type *CCC* through x (which are precisely the lines through x which do not intersect \mathcal{S}) is equal to $|Y|/2 = 8$. Hence, there are 9 lines through x , exactly one of which, the line of type *ACC*, intersects \mathcal{S} .

- (2) If $y_1, y_2 \in Y$ with $y_1 \neq y_2$, then $f_{y_1} \neq f_{y_2}$, since $\mathcal{O}_{f_{y_1}} \neq \mathcal{O}_{f_{y_2}}$. Therefore, all points of the 8 lines of type *CCC* through x induce distinct valuations of type *C*. This gives rise to 8 distinct \mathcal{V} -lines of type *CCC* through the valuation f_x . By Table 2.4, these are all the type *CCC* lines of \mathcal{V} through f_x .
- (3) In view of Lemma 4.2.5, it suffices to prove that for any two type *C* valuations f_1 and f_2 which are collinear in \mathcal{V} , the number of times f_1 is induced by a point of \mathcal{S}' is equal to the number of times f_2 is induced. This follows directly from (2), as for

every point x for which $f_x = f_1$, there is a unique point y for which $f_y = f_2$, and vice versa. □

Corollary 4.2.7. *Every generalized hexagon containing $\mathcal{S} \cong H(2)^D$ as a full subgeometry is finite.*

Proof. Let \mathcal{S}' be such a generalized hexagon. Then by Lemma 1.2.6, \mathcal{S}' is also thick, and hence has an order. Therefore, it suffices to show that there exists a point in \mathcal{S}' through which there are only finitely many lines. By Lemma 1.2.5, \mathcal{S} is isometrically embedded in \mathcal{S}' . Since there are no quads in a generalized hexagon, we see that every point of $\Gamma_1(\mathcal{S})$ has exactly nine lines through it, and thus \mathcal{S}' is finite. □

Lemma 4.2.6(3) shows that the map ψ when restricted to the points of \mathcal{S}' is a surjection between the set of points of \mathcal{S}' and the valuations of \mathcal{S} that have type A or C . Moreover, if there are no quads intersecting \mathcal{S} , then the preimage of each valuation under this map has exactly k elements for some constant k . Clearly this condition is satisfied if \mathcal{S}' is a generalized hexagon, which does not have any quads. We will now show that in general \mathcal{S}' does not have any quads that intersect \mathcal{S} , from which we will later deduce that \mathcal{S}' is a generalized hexagon.

We already know of one generalized hexagon which contains $\mathcal{S} \cong H(2)^D$ as a full proper subgeometry, the dual twisted triality hexagon $T(2, 8)$. From Proposition 1.2.4, $T(2, 8)$ has $819 = 63 + 3 \cdot 252$ points. Since there are precisely 252 type C valuations in \mathcal{V} , from the discussion so far we deduce that every type C valuation is induced exactly 3 times in $T(2, 8)$. This embedding of $H(2)^D$ in $T(2, 8)$ can now further help us in obtaining certain properties of the valuation geometry \mathcal{V} which will be useful to us. These properties can alternatively be checked in the computer model of \mathcal{V} .

Lemma 4.2.8. *Let \mathcal{V}' be the subgeometry of \mathcal{V} obtained by taking only points of type C and the lines of type CCC . Then the following holds.*

- (1) *If f_1, f_2 are distinct collinear points of \mathcal{V}' , then the unique points in \mathcal{O}_{f_1} and \mathcal{O}_{f_2} are at distance 3 from each other.*
- (2) *Suppose G is a (3×3) -subgrid of \mathcal{V}' . Let f_1, f_2 be two noncollinear points of G and for $i \in \{1, 2\}$, let x_i denote the unique point of \mathcal{S} in \mathcal{O}_{f_i} . Then x_1, x_2 are distinct and $d(x_1, x_2) = 3$.*

Proof. Just for the scope of this proof, let $\mathcal{S}' \cong T(2, 8)$.

- (1) Let f_1, f_2 be two distinct collinear points of \mathcal{V}' . We know that the map ψ in this case is surjective, and therefore we can pick two collinear type C points x_1, x_2 in \mathcal{S}' such that $f_1 = f_{x_1}$ and $f_2 = f_{x_2}$. Let x'_1 and x'_2 be the unique points of $\mathcal{S} \cong H(2)^D$ collinear with x_1 and x_2 , respectively. Then we have $\{x'_1\} = \mathcal{O}_{f_1}$ and $\{x'_2\} = \mathcal{O}_{f_2}$. The distance between x'_1 and x'_2 must be at least 3 as otherwise we will get an ordinary m -gon for $m < 6$. Therefore, $d(x'_1, x'_2) = 3$.
- (2) The fact that $d(x_1, x_2) \geq 2$ follows from a similar reasoning as (1) inside $\mathcal{S}' \cong T(2, 8)$. To prove that the points are in fact at distance 3 from each other, we use the computer model of \mathcal{V} as we cannot find an easy argument using $T(2, 8)$ which shows this. (Computationally we can see that if f_1 and f_2 are any two non-collinear points of \mathcal{V}' which have a common neighbor, then it is not necessarily true that $d(x_1, x_2) = 3$.)

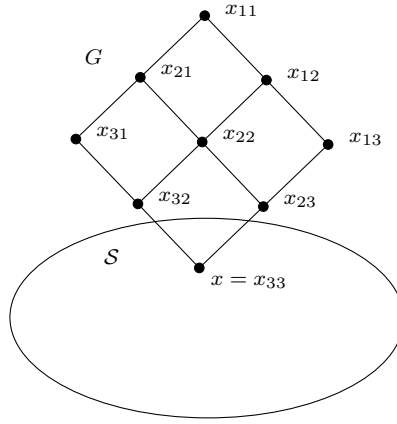


Figure 4.1.: Lemma 4.2.9 (5)

□

Lemma 4.2.9. *There are no quads in \mathcal{S}' which contain a point of \mathcal{S} .*

Proof. Let Q be a quad in \mathcal{S}' which contains a point of \mathcal{S} . Since Q has three points on each line, we can find a (3×3) -subgrid G which also contains a point of \mathcal{S} . Therefore we have two subspaces of \mathcal{S}' , the grid G and the subhexagon \mathcal{S} , which intersect non-trivially. The intersection $G \cap \mathcal{S}$ must also be a subspace and thus we have the following cases.

- (1) The grid G is completely contained in \mathcal{S} , which is impossible since \mathcal{S} is a generalized hexagon.
- (2) G intersects \mathcal{S} in a union $L_1 \cup L_2$ of two intersecting lines. But then, any point of $G \setminus (L_1 \cup L_2)$, which is necessarily in $\Gamma_1(\mathcal{S})$ and of type C , will be collinear with two points of \mathcal{S} , a contradiction to Lemma 4.2.2.
- (3) G intersects \mathcal{S} in 2 or 3 mutually non-collinear points. Then take a common neighbor of any two points in $G \cap \mathcal{S}$. Just like in (2), this gives us a contradiction as this point is collinear with at least two points of \mathcal{S} .
- (4) G intersects \mathcal{S} in a line L . Take a line L' of G which is disjoint from L . Take two points x_1 and x_2 on L , necessarily of type C , and let y_1 and y_2 be the unique points of L collinear with x_1 and x_2 , respectively. Then f_{x_1} and f_{x_2} are two distinct type C valuations of \mathcal{S} for which the unique points in $\mathcal{O}_{f_{x_1}}$ and $\mathcal{O}_{f_{x_2}}$ are collinear, contradicting Lemma 4.2.8(1).
- (5) G intersects \mathcal{S} in a point x . We label the points of G by x_{ij} , $i, j \in \{1, 2, 3\}$, such that $x = x_{33}$ and x_{ij} is collinear with $x_{i'j'}$ if and only if either $i = i'$ or $j = j'$ (see Figure 4.1). Let f_{ij} denote the valuation induced by the point x_{ij} of the grid. Then by a direct reasoning in $T(2, 8)$, we must have $f_{31} = f_{23}$ and $f_{31} = f_{13}$. Which means that in the line $\{x_{13}, x_{23}, x_{33}\}$ of \mathcal{S}' , we have $f_{13} = f_{23}$, contradicting Theorem 2.3.2.

□

Corollary 4.2.10. *Every point of \mathcal{S}' not contained in \mathcal{S} is incident with precisely nine lines.*

Proof. This follows from Lemmas 4.2.6 and 4.2.9. □

One way to proceed now would be to directly show that there are no quads in \mathcal{S}' , but this seems to be difficult. We proceed indirectly by first proving that every point of \mathcal{S}' is incident with precisely 9 lines.

Lemma 4.2.11. *Let x be a point of \mathcal{S}' not contained in \mathcal{S} , let x' be the unique point of \mathcal{S} collinear with x and let y be a point of \mathcal{S} which is at distance 2 from x' . Then $d(x, y) = 3$ and every neighbor of y has at most one common neighbor with x .*

Proof. Let u denote the unique common neighbor of x' and y , which necessarily lies in \mathcal{S} by Lemma 4.2.2(3). Then $d(x, u) = 2$. If $d(x, y) \neq 3$, i.e., $d(x, y) \leq 2$, then **NP2** would imply that the line uy contains a point at distance 1 from x , contradicting the fact that \mathcal{O}_{f_x} is a singleton (see again Lemma 4.2.2(3)).

Now let z be a neighbor of y such that x and z have more than one common neighbor. Then by Theorem 1.3.8 there exists a quad Q , and thus a (3×3) -subgrid G , through x and z . By Lemma 4.2.9, we have $Q \cap \mathcal{S} = G \cap \mathcal{S} = \emptyset$, and in particular z lies outside \mathcal{S} . If two points of Q induce the same valuation, then they would have a common neighbor in \mathcal{S} which by the convexity of Q would lie in Q , thus contradicting Lemma 4.2.9. Therefore, each of the 9 points in G induce distinct type C valuations, and hence G induces a grid G' in the subgeometry \mathcal{V}' of \mathcal{V} formed by taking type C points and type CCC lines. Since x, z are noncollinear in G , f_x, f_z are noncollinear in \mathcal{V}' . We now get a contradiction to Lemma 4.2.8(2) because $\mathcal{O}_{f_x} = \{x'\}$, $\mathcal{O}_{f_z} = \{y\}$ and $d(x', y) = 2$. \square

Lemma 4.2.12. *Every point of \mathcal{S} is incident with exactly 9 lines in \mathcal{S} , and hence \mathcal{S}' has order $(2, 8)$.*

Proof. Let x be a point of \mathcal{S} and y another point of \mathcal{S} at distance 2 from x . Let f be a valuation of type C for which $\mathcal{O}_f = \{y\}$ (such a valuation always exists). Then by Lemma 4.2.6(3), there exists a point z in \mathcal{S}' such that $f_z = f$. This point z is collinear with y . By Lemma 4.2.11, $d(z, x) = 3$. We will show that there is a bijection between the lines through z (there are exactly 9 of them by Lemma 4.2.6) and the lines through x .

We count the number of paths z, u, v, x of length 3 between z and x . On each of the 9 lines through z , by **NP2**, there exists a unique point of $\Gamma_2(x)$, and therefore there are 9 possibilities for u . Since there are no quads through x , every point of $\Gamma_1(z) \cap \Gamma_2(x)$ has a unique common neighbor with x . Therefore, for each of the 9 possibilities for u , there is a unique possibility for v , giving rise to exactly 9 such paths. Now let $t_x + 1$ be the (possibly infinite) number of lines through x . Each such line contains a unique point at distance 2 from z , giving rise to $t_x + 1$ choices for v . By Lemma 4.2.11, for each such v there is a unique u (which is a common neighbor of z and v). Therefore the number of such paths is $t_x + 1$, and hence $t_x + 1 = 9$. \square

Now Theorem 4.2.1(2) can be proved as follows. The total number of points in \mathcal{S}' is $|\mathcal{S}| + 12 \cdot |\mathcal{S}| = 63 + 12 \cdot 63 = 819$ since each point of \mathcal{S} is incident with precisely 6 lines of type ACC , each of which gives rise to 2 points of \mathcal{S}' not contained in \mathcal{S} and this covers all points of \mathcal{S}' by Lemma 4.2.2. It now follows from Lemma 1.3.6 that \mathcal{S}' is a generalized hexagon of order $(2, 8)$ as $819 = (1 + 2)(1 + 2 \cdot 8 + 2^2 \cdot 8^2)$. This generalized hexagon is necessarily isomorphic to $T(2, 8)$ by [53].

4.2.2. Near Hexagons Containing $H(2)$

For this section, let \mathcal{S}' be a near hexagon with three points on each line which contains $\mathcal{S} \cong H(2)$ isometrically embedded in it as a full subgeometry. We will use the same convention as before in assigning types to the points and lines of \mathcal{S}' based on the types in the valuation geometry of $H(2)$ (see Tables 2.5 and 2.6). Again from the basic theory of valuations we see that every point of \mathcal{S}' has one of the 7 types mentioned in Table 2.5 and every line has one of the 20 types mentioned in Table 2.6. Note that the type C valuation is *ovoidal*.

Lemma 4.2.13. (1) *Every point of \mathcal{S}' is at distance at most 2 from \mathcal{S} .*

(2) *Every point x of \mathcal{S} has type A , and the valuation f_x is classical with center x .*

(3) *Every point of \mathcal{S}' at distance 1 from \mathcal{S} has type B_i for some $i \in \{1, \dots, 5\}$.*

(4) *Every point of \mathcal{S}' at distance 2 from \mathcal{S} has type C .*

Proof. The proof is similar to Lemma 4.2.2, with the only complication occurring in the case when there is a point x at distance 1 inducing a type C (ovoidal) valuation. So let x be such a point, and let $O = \mathcal{O}_{f_x}$ be the ovoid of \mathcal{S} induced by points of \mathcal{S} that are collinear with x . We show that there exist two points in O which are at distance 3 from each other, thus contradicting the assumption that \mathcal{S} is isometrically embedded in \mathcal{S}' .

Let $y, z \in O$ such that $d(y, z) = 2$ (if no such pair exists, then there is nothing to prove). Let w be their common neighbor in \mathcal{S} . Since O is a 1-ovoid, $w \notin O$. Now let w' be the third point on the line wz , and L any line through w' other than the line wz . Then $w' \notin O$, and every point in $L \setminus \{w'\}$ has distance 3 from y because \mathcal{S} is a generalized hexagon. At least one of the points in $L \setminus \{w'\}$ must be contained in O , giving us two points of O at distance 3 from each other. \square

We can now prove (without invoking the classification result of [53]) that if \mathcal{S}' is a generalized hexagon then it must be equal to \mathcal{S} , and in particular finite. We will give another proof of this result in Section 4.3.1.

Theorem 4.2.14. *There exists no generalized hexagon which contains $H(2)$ as a full proper subgeometry.*

Proof. Let \mathcal{S}' be a generalized hexagon containing $\mathcal{S} \cong H(2)$ as a full subgeometry. By Lemma 1.2.6, it has an order $(2, t)$ where t is a possibly infinite cardinal. We shall show that the only possible valuations of \mathcal{S} that points of \mathcal{S}' can induce are the type A (classical) valuations, which by Lemma 4.2.13 would imply that $\mathcal{S}' = \mathcal{S}$. We do so in the following sequence of steps.

(1) *There is no point in \mathcal{S}' of type B_i with $i > 1$.* Let x be such a hypothetical point. By Lemma 4.2.13, x must be at distance 1 from \mathcal{S} . Since $|\mathcal{O}_{f_x}| \geq 2$ (see Table 2.5) we can take two distinct points y_1, y_2 in \mathcal{O}_{f_x} . These points are at distance 2 in \mathcal{S}' as x is their common neighbor, and hence at distance 2 in \mathcal{S} because \mathcal{S} is isometrically embedded. Therefore, they have a common neighbor $y \neq x$ in \mathcal{S} . This contradicts the fact that every pair of points at distance 2 in a generalized hexagon have a unique common neighbor.

- (2) *There is no point in \mathcal{S}' of type B_1 .* Let x be a point of type B_1 in \mathcal{S}' . Then by Lemma 4.2.13, x is at distance 1 from \mathcal{S} and since \mathcal{O}_{f_x} is a singleton, there is a unique point $\pi(x)$ in \mathcal{S} collinear with x . The valuation f_x has 22 points of value 1 (by Table 2.5) which by Theorem 2.2.3 is equivalent to the fact that there are exactly 22 points in \mathcal{S} which are at distance 2 from x . Out of these 22 points, 6 are neighbors of $\pi(x)$ in \mathcal{S} and therefore each of the remaining 16 must have a common neighbor with x which lies outside \mathcal{S} . Since x is collinear with only type B_1 points in $\Gamma_1(\mathcal{S})$ (as by (1), there are no such points of type B_i for $i > 2$ and by Lemma 4.2.13 there are no type C points in $\Gamma_1(\mathcal{S})$), all of which have a unique neighbor in \mathcal{S} , we get at least 16 points of type B_1 in $\Gamma_1(\mathcal{S})$ collinear with x . All of these neighbors of x must induce distinct valuations since their zero sets (\mathcal{O}_f) are distinct. Therefore, we get 16 type B_1 neighbors of the valuation f_x in the valuation geometry \mathcal{V} , which contradicts the entry in row $B_1 B_1 B_1$ and column B_1 of Table 2.6 which says that there are only 6 such neighbors.
- (3) *There is no point in \mathcal{S}' of type C .* By Table 2.6, we see that every line through such a point must contain a point of type B_i for some $i > 1$, but no such point exists by (1).

□

We now prove that every *near hexagon* \mathcal{S}' containing $\mathcal{S} \cong H(2)$ as a full isometrically embedded subgeometry is finite. The reasoning below gives an alternate proof of Theorem 4.2.14 if we invoke the classification of finite generalized hexagons with three points on each line [53] and the observation that none of the hexagons appearing in the classification contain $H(2)$ as a proper subgeometry. One of the main results we use in these proofs is Lemma 1.3.9 which in particular shows that every pair of points in \mathcal{S}' at distance 2 from each other have finitely many common neighbors.

Lemma 4.2.15. *For every $i \in \{2, 3, 4, 5\}$ there are only finitely many points in \mathcal{S}' of type B_i .*

Proof. Let x be a point of type B_i for some $i > 1$. From Table 2.5 there are at least two distinct points y and z of \mathcal{S} in \mathcal{O}_{f_x} , necessarily collinear with x . By **NP2** y and z must be at distance 2 from each other. Now, y and z have at most five common neighbors (Lemma 1.3.9) and one of these must be contained in \mathcal{S} . From this it follows that the number of points of type B_i for some $i \in \{2, 3, 4, 5\}$ is at most 4 times the number of unordered pairs $\{p, q\}$ of points of \mathcal{S} at distance 2 from each other, which is finite (in fact equal to 3024). □

Lemma 4.2.16. *There are only finitely many points of type B_1 in \mathcal{S}' .*

Proof. Let \mathcal{B} denote the set of those points of \mathcal{S}' that have type B_i for some $i \in \{2, 3, 4, 5\}$. Then \mathcal{B} is finite by Lemma 4.2.15. Let \mathcal{A} denote the set of those points of \mathcal{S}' that have type A , i.e., the points of \mathcal{S} . Then the set $\mathcal{A} \cup \mathcal{B}$ is also finite. Let x be a point of type B_1 in \mathcal{S} . Then by Lemma 4.2.13, x is at distance 1 from \mathcal{S} , and since \mathcal{O}_{f_x} is a singleton, there exists a unique point $\pi(x)$ in \mathcal{S} collinear with x . If x is only collinear with points of type A , B_1 or C , then by the same reasoning as in the proof of Theorem 4.2.14, we get a contradiction. So, x is collinear with at least one point of \mathcal{B} , and we have already seen that it is collinear with at least one point of \mathcal{A} . Thus we see that x is the common neighbor of two points at distance 2 in the finite set $\mathcal{A} \cup \mathcal{B}$. Since each such pair of points at distance 2 in the near polygon \mathcal{S}' has finitely many (at most 5) common neighbors, we

see that the set of points of type B_1 must be finite; in fact, the cardinality of this set is bounded above by 5 times the number of pairs of points at distance 2 in $\mathcal{A} \cup \mathcal{B}$. \square

Lemma 4.2.17. *There are only finitely many points of type C in \mathcal{S}' .*

Proof. Let x be a point of type C in \mathcal{S}' . Then the set $\Gamma_2(x) \cap \mathcal{S}$ is a 1-ovoid of \mathcal{S} , and hence it has cardinality 21. Let S_x be the set of common neighbors between x and elements of \mathcal{O}_{f_x} (the 1-ovoid of \mathcal{S} induced by x). By Lemma 4.2.13, each element y of S_x has type B_i for some $i \in \{1, \dots, 5\}$ and hence by Table 2.5 y is collinear with at most 9 points of \mathcal{S} . Therefore $|S_x| \geq 2$, and we get two points of the set $\Gamma_1(\mathcal{S})$ at distance 2 from each other having x as a common neighbor. By Lemmas 4.2.15 and 4.2.16, the set $\Gamma_1(\mathcal{S})$ is finite, and thus by a same reasoning as before there are only finitely many points in \mathcal{S}' of type C . \square

From Lemmas 4.2.15, 4.2.16 and 4.2.17 it follows that \mathcal{S}' must be finite. We finally note that there are near hexagons which are not generalized hexagons and contain $\mathcal{S} \cong \text{H}(2)$ as a full isometrically embedded subgeometry, for example, the dual polar spaces $\text{DW}(5, 2)$ and $\text{DH}(5, 4)$.¹ We do not provide any classification of these near hexagons.

4.3. Polygonal Valuations and Intersections of Hyperplanes

While the theory of valuations outlined in Chapter 2 works quite well for near polygons with three points on each line, for other near polygons it is difficult in practice. Firstly, in the case when there are more than three points on a line, we do not have a good way of computing hyperplanes; and thus the approach outlined in Section 3.4 does not work. So, we might have to stick to a direct backtrack search to find all the valuations, which will only work for really small cases. Second, we lose the power of Theorem 2.3.2 which lets us define the valuation geometry by simply taking pairs of distinct neighboring valuations, instead of arbitrary triples of pairwise neighboring valuations.

In this section, we will give a short summary of the theory of so-called *polygonal valuations* (developed by De Bruyn in [60]) which is more restrictive than the one specified in Chapter 2 and only suitable to study generalized polygons containing a subpolygon. This theory will be used to derive some counting based results on intersections of hyperplanes in arbitrary generalized hexagons with an order (Lemma 4.3.7), which will help us prove some results for generalized hexagons containing $\text{H}(3)$, $\text{H}(4)$ or $\text{H}(4)^D$ as a full subgeometry (Theorem 4.3.11). Note that in contrast with Section 4.2 we will not be able to say anything about near hexagons which are not generalized hexagons and contain these subhexagons.

Definition. Suppose $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \text{I})$ is a generalized $2n$ -gon with $n \in \mathbb{N} \setminus \{0, 1\}$. Then a *polygonal valuation* of \mathcal{S} is a map $f : \mathcal{P} \rightarrow \mathbb{N}$ that satisfies the following axioms:

- (PV1) There exists at least one point with f -value 0.
- (PV2) Every line L of \mathcal{S} contains a unique point x_L such that $f(x) = f(x_L) + 1$ for all points $x \neq x_L$ incident with L .

¹In fact, this kind of explains why the valuation geometry of $\text{H}(2)$ is much larger than that of $\text{H}(2)^D$.

(PV3) Let M_f denote the maximum value of f over the points of \mathcal{S} . If x is a point with $f(x) < M_f$, then there is at most one line through x containing a (necessarily unique) point y satisfying $f(y) = f(x) - 1$.

Or in other words, a polygonal valuation is a valuation of \mathcal{S} which satisfies the condition (PV3). It turns out that with this one extra condition, we have a much stronger theory where several interesting and new results hold. For example, from the following result it follows that the polygonal valuations of a generalized polygon are in bijective correspondence with its hyperplanes of valuation type.

Lemma 4.3.1 ([60, Propostion 3.10]). *If f is a polygonal valuation of a generalized $2d$ -gon $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$, then $f(x) = M_f - d(x, \mathcal{P} \setminus H_f)$ for every point x of \mathcal{S} , where H_f is the hyperplane of \mathcal{S} corresponding to f .*

It is easy to check that the classical, semi-classical and ovoidal valuations defined in Chapter 2, all satisfy (PV3). And hence, these are all examples of polygonal valuations as well. In fact, for generalized hexagons these are the only possible polygonal valuations, as the following lemma shows.

Lemma 4.3.2 ([60, Propostion 3.3]). *Suppose f is a polygonal valuation of a generalized $2n$ -gon \mathcal{S} . Then:*

- (1) *f is an ovoidal polygonal valuation if and only if $M_f = 1$;*
- (2) *f is a semi-classical polygonal valuation if and only if $M_f = n - 1$;*
- (3) *f is a classical valuation if and only if $M_f = n$.*

Corollary 4.3.3. *If \mathcal{S} is a generalized hexagon, then every polygonal valuation of \mathcal{S} is classical, semi-classical or ovoidal.*

Proof. As noted in Section 2.2, the maximum value of a valuation (and hence a polygonal valuation) is bounded above by the diameter of the near polygon. Therefore, for any polygonal valuation f of a generalized hexagon, we have $M_f \in \{1, 2, 3\}$. \square

A quick glance at Table 2.3 of Chapter 3 will show that Corollary 4.3.3 does not hold for valuations of near hexagons, as the type B and type D valuations are not classical, semi-classical or ovoidal. Since we know how to compute these three types of valuations, we can use the machinery of Chapter 3 to find out all polygonal valuations of generalized hexagons (even those which have more than three points on each line). We now state the analogous result of Theorem 2.2.3 for polygonal valuations.

Lemma 4.3.4 ([60, Proposition 6.1]). *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a generalized $2n$ -gon contained in a generalized $2n$ -gon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathbf{I}')$ as a full subgeometry. Let x be a point of \mathcal{S}' and put $m := d(x, \mathcal{P})$. Noting that $m \in \{0, 1, \dots, n - 1\}$, we define $f_x(y) := d(x, y) - m$ for every point $y \in \mathcal{P}$. Then:*

- (1) *f_x is a valuation of \mathcal{S} with $M_{f_x} = n - m$.*
- (2) *The valuation f_x is classical if and only if x is a point of \mathcal{S} , semi-classical if and only if $m = 1$ and ovoidal if and only if $m = n - 1$.*
- (3) *If x_1 and x_2 are two distinct collinear points of \mathcal{S} , then the valuations f_{x_1} and f_{x_2} are distinct.*

For a result analogous to Theorem 2.2.6, we define the notion of *L-sets* and *admissible sets* of polygonal valuations of a generalized $2n$ -gon \mathcal{S} , which will correspond to the lines of a generalized $2n$ -gons containing \mathcal{S} as a full subgeometry. This notion will also allow us to define the notion of *polygonal valuation geometry*, that is useful in classifying generalized polygons containing a fixed subpolygon.

Definition. Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a generalized $2n$ -gon. A collection $\mathcal{F} = \{f_i : i \in I\}$ of distinct polygonal valuations of \mathcal{S} , where I is an index set of cardinality at least 2, is called an *L-set* if for every $x \in \mathcal{P}$, there exists an $i \in I$ such that $f_j(x) - M_{f_j} = f_i(x) - M_{f_i} + 1$ for every $j \in I \setminus \{i\}$. The set \mathcal{F} is called *admissible* if the following holds for all $i, j \in I$ with $i \neq j$, for all $x \in \mathcal{P}$ such that x is not collinear with any point with f_i -value $f_i(x) - 1$, and for all $y \in \mathcal{P}$ such that y is not collinear with any point with f_j -value $f_j(y) - 1$:

- (1) if f_i and f_j are classical, then $d(x, y) = 1$;
- (2) if $x = y$, then $(f_i(x) - M_{f_i}) - (f_j(x) - M_{f_j}) \in \{-1, 0, 1\}$;
- (3) if $x \neq y$ and at least one of f_i, f_j is not classical, then $d(x, y) + f_i(x) + f_j(y) - M_{f_i} - M_{f_j} + 1 \geq 0$.

Lemma 4.3.5 ([60, Proposition 6.2]). *Let \mathcal{S} be a generalized $2n$ -gon contained in a generalized $2n$ -gon \mathcal{S}' as a full subgeometry. For every line L of \mathcal{S}' define $\mathcal{F}_L = \{f_x : x \in L\}$, where f_x denotes the valuation of \mathcal{S} induced by the point x . Then \mathcal{F}_L is an admissible L-set of polygonal valuations.*

Definition. The *polygonal valuation geometry* of a generalized $2n$ -gon \mathcal{S} is the point-line geometry $\mathcal{V}_{\mathcal{S}}$ whose points are the polygonal valuations of \mathcal{S} and lines are the admissible L-sets of polygonal valuations of \mathcal{S} , with incidence as set containment.

Just like the notion of valuation geometry defined in Chapter 2 for near polygons with three points on each line, this notion of polygonal valuation geometry for generalized polygons (with any number of lines per point) can help us study generalized polygons containing a fixed generalized polygon as a full subgeometry. This is precisely what was used by De Bruyn in [62] to prove that there is a unique generalized octagon of order $(2, 4)$ containing a suboctagon of order $(2, 1)$. In fact, polygonal valuation geometry of the split Cayley hexagon $H(3, 1)$ (which is not too hard to compute using the tools described in Chapter 3) can be used to prove the result of De Medts and Van Maldeghem [67] that $H(3)$ is the unique generalized hexagon containing a subhexagon of order $(3, 1)$. This is an unpublished proof by De Bruyn which was mentioned in [61]. For our purposes, we do not need the full power of polygonal valuation geometry, but only the following result on hyperplanes that follows directly from Lemma 4.3.5 and [60, Proposition 4.7], and its consequences.

Lemma 4.3.6. *Let $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathbf{I})$ be a generalized $2n$ -gon contained in a generalized $2n$ -gon $\mathcal{S}' = (\mathcal{P}', \mathcal{L}', \mathbf{I}')$ as a full subgeometry, and let L be a line of \mathcal{S}' . For a point x of \mathcal{S}' , let H_x denote the hyperplane of \mathcal{S} formed by points of \mathcal{S} at distance at most $n - 1$ from x . Then:*

- (1) *the set of hyperplanes $\{H_x : x \in \mathcal{P}', x \notin L\}$ covers \mathcal{P} ;*
- (2) *if x_1, x_2 and x_3 are three distinct points on L , then $H_{x_1} \cap H_{x_2} = H_{x_1} \cap H_{x_3}$.*

A finite generalized hexagon of order (s, t) has $(1+s)(1+st+s^2t^2)$ points. For each of the three types of hyperplanes in generalized hexagons corresponding to the three types of polygonal valuations, we can determine the sizes of the hyperplanes by simple counting. The singular, semi-singular and ovoidal hyperplanes in a finite generalized hexagon of order (s, t) are of sizes $1 + s(t+1) + s^2t(t+1)$, $1 + s(t+1) + s^2t^2$ and $1 + st + s^2t^2$ respectively. The following is the main result on hyperplane intersections that we will use to prove Theorem 4.3.11.

Lemma 4.3.7. *Let \mathcal{S} be a finite generalized hexagon of order (s, t) contained in a generalized hexagon \mathcal{S}' as a full subgeometry, and let L be a line of \mathcal{S}' that does not intersect \mathcal{S} . Let n_L denote the number of points on L that are at distance 2 from \mathcal{S} . For a point x in \mathcal{S}' , let H_x denote the hyperplane of \mathcal{S} formed by taking points of \mathcal{S} that are at non-maximal distance from x . Then for any two distinct points x and y on L we have $|H_x \cap H_y| = s + 1 - n_L$.*

Proof. By Lemma 1.2.6 and the fact that opposite lines are incident with the same number of points, we know that every line of \mathcal{S}' is incident with precisely $s + 1$ points. By Lemma 4.3.4, for every point x on L the hyperplane H_x of \mathcal{S} is either semi-singular or ovoidal. The points on L that are at distance 1 from \mathcal{S} induce semi-singular hyperplanes, while those at distance 2 induce ovoidal hyperplanes. By Lemma 4.3.6 there exists a fixed subset X of points of \mathcal{S} such that $H_x \cap H_y = X$ for every pair of distinct points x, y on L , and every point of \mathcal{S} is contained in some hyperplane induced by a point on L . Let the size of X be k . There are n_L hyperplanes of size $1 + st + s^2t^2$ (ovoidal) and $s + 1 - n_L$ hyperplanes of size $1 + s(t+1) + s^2t^2$ (semi-singular) which cover a set of size $(1+s)(1+st+s^2t^2)$ (points of \mathcal{S}) and pairwise intersect in k points. Therefore, we have

$$n_L(1 + st + s^2t^2 - k) + (s + 1 - n_L)(1 + s(t+1) + s^2t^2 - k) + k = (1+s)(1+st+s^2t^2),$$

which can be solved for k to get $k = s + 1 - n_L$. \square

Lemma 4.3.8. *Let \mathcal{S} be a finite generalized hexagon of order (s, t) having the property that $|H_1 \cap H_2| > s + 1$ for any two semi-singular hyperplanes H_1 and H_2 of \mathcal{S} whose centers lie at distance 3 from each other. Then there does not exist any generalized hexagon that contains \mathcal{S} as a full proper subgeometry.*

Proof. Say there is such a generalized hexagon \mathcal{S}' and let x be a point of \mathcal{S}' that is at distance 1 from the point set of \mathcal{S} . By Lemma 4.3.4, H_x is a semi-singular hyperplane, corresponding to the semi-classical polygonal valuation f_x defined by $f_x(y) = d(x, y) - 1$ for points y of \mathcal{S} . Let x' be the unique point of \mathcal{S} with f_x -value 0, or equivalently the unique point of \mathcal{S} at distance 1 from x . Let O' be the 1-ovoid in the subgeometry of \mathcal{S} induced on $\Gamma_3(x')$ which defines the hyperplane H_x . Let y be a point of O' . Then $f_x(y) = 1$, and hence $d(x, y) = 2$. Let z be a common neighbor of x and y . Then z must lie outside \mathcal{S} , and the line $L = xz$ does not contain any point of \mathcal{S} . Note that f_z is also a semi-classical polygonal valuation since z has distance 1 from \mathcal{S} . From Lemma 4.3.7 it follows that $|H_x \cap H_z| \leq s + 1$. Moreover we have $d(x', y) = 3$, thus contradicting the assumption stated in the lemma. \square

We will use Lemma 4.3.8 to prove Theorem 4.3.11(2) in Section 4.3.1. The following result will help us prove the finiteness of generalized hexagons containing a subhexagon.

Lemma 4.3.9. *Let \mathcal{S} be a finite generalized hexagon with only thick lines that is contained in a generalized hexagon \mathcal{S}' as a full subgeometry. If every point of \mathcal{S}' is at distance at most 1 from \mathcal{S} , then \mathcal{S}' is also finite.*

Proof. Lemma 1.2.6 implies that every line of \mathcal{S}' is also thick and hence has an order. Suppose now that every point of \mathcal{S}' is at distance at most 1 from \mathcal{S} and that $\mathcal{S}' \neq \mathcal{S}$. Let x be a point in \mathcal{S}' at distance 1 from \mathcal{S} . Then it suffices to show that there are only finitely many lines through x .

Note that x induces a semi-classical polygonal valuation on \mathcal{S} , and thus there exists a unique point y of \mathcal{S} with f_x -value 0, which by Lemma 4.3.4 is the unique point of \mathcal{S} at distance 1 from x . Therefore, there is a unique line through x in \mathcal{S}' which meets \mathcal{S} . Now, let L be any other line through x . Pick any point z in $L \setminus \{x\}$. Then z is again collinear with a unique point z' of \mathcal{S} as every point of \mathcal{S}' , and in particular z , is at distance at most 1 from \mathcal{S} . In this manner we can associate each line of \mathcal{S}' through x that does not intersect \mathcal{S} with a point of \mathcal{S} . Moreover, for two distinct lines L_1, L_2 through x not meeting \mathcal{S} the points z'_1, z'_2 of \mathcal{S} obtained in this manner by taking points $z_1 \in L_1 \setminus \{x\}$ and $z_2 \in L_2 \setminus \{x\}$ must be distinct, as otherwise we will get a pair of points at distance 2 from each in the generalized hexagon \mathcal{S}' that have at least two common neighbors. Since the number of points in \mathcal{S} is finite, this shows that there are only finitely many lines through x . \square

Corollary 4.3.10. *If a generalized hexagon \mathcal{S} does not have any 1-ovoids, then it cannot be contained in a semi-finite generalized hexagon as a full subgeometry.*

Proof. Let \mathcal{S}' be a generalized hexagon containing \mathcal{S} as a full subgeometry. Say \mathcal{S} does not have any 1-ovoid. Then by Lemma 4.3.4 every point of \mathcal{S}' is at distance at most 1 from \mathcal{S} , and so \mathcal{S}' should be finite by Lemma 4.3.9. \square

Remark. In [66], De Bruyn and Vanhove showed that no generalized hexagon of order (s, s^3) can have 1-ovoids, thus proving that these hexagons cannot be contained in a semi-finite generalized hexagon as a full subgeometry (see Corollary 3.20 in [66]).

4.3.1. Generalized Hexagons Containing Split Cayley Hexagons and their Dual

We give a computer-aided proof of the following result in this section.

Theorem 4.3.11. *Let $q \in \{2, 3, 4\}$ and let \mathcal{S} be a generalized hexagon isomorphic to the split Cayley hexagon $H(q)$ or its dual $H(q)^D$. Then the following holds for any generalized hexagon \mathcal{S}' that contains \mathcal{S} as a full subgeometry:*

- (1) \mathcal{S}' is finite;
- (2) if $q \in \{2, 4\}$ and $\mathcal{S} \cong H(q)$, then $\mathcal{S}' = \mathcal{S}$.

From Corollary 4.3.10 and the non-existence of 1-ovoids in $H(4)^D$ proved in Section 3.5 it follows that every generalized hexagon containing a subhexagon isomorphic to $H(4)^D$ is finite. The same holds for generalized hexagons containing $H(2)^D$ as by Table 2.3, it does not have any 1-ovoids.

Now let $\mathcal{S} \cong H(3)$ and let \mathcal{S}' be a generalized hexagon containing $H(3)$ as a full subgeometry. Say there exists a point x in \mathcal{S}' at distance 2 from \mathcal{S} and let x, y, z be a path of length 2 from x to a point z of \mathcal{S} . Then by Lemma 4.3.7 we have $|H_x \cap H_y| \leq 3$, where H_x is the 1-ovoid of \mathcal{S} induced by x and H_y is the semi-singular hyperplane of \mathcal{S} induced by y . Note that $z \in H_x$ and z is the center of H_y . Since the geometry is small enough and

the automorphism group acts transitively on the points, we can fix a point p of $H(3)$ in our computer model obtained in Chapter 3 and look at the intersection sizes of 1-ovoids through p and semi-singular hyperplanes with center p . The following SageMath code can be used for this computation. We assume that the SageMath function `ovoids` defined in Chapter 3 is available to us, and that the function `get_data` gives us the points, lines and the distance matrix of a point-line geometry (which can be implemented using the methods of Chapter 3).

```
def semi_singular_hyperplanes(points, lines, D, p):
    """
    Computes all semi-singular hyperplanes of a point line
    geometry with a fixed point as center.

    Args:
    points -- a list of points
    lines -- a list of lines
    D -- the distance matrix of the collinearity graph
    p -- the fixed point

    Returns:
    a list of all semi-singular hyperplanes with center p
    """
    dist3 = [x for x in points if D[x][p] == 3]
    L = [l for l in lines if any([x in dist3 for x in l])]
    L = [[x for x in l if x in dist3] for l in L]
    ov = list(ovoids(dist3, L))
    return [sorted([x for x in points if x in O or D[x][p] <= 1]) for O in ov]

# H(3)
points, lines, D = get_data("H3")
p = points[0]
ovoids_p = [O for O in ovoids(points, lines) if p in O]
hyp_p = semi_singular_hyperplanes(points, lines, D, p)
S = {len(set(x) & set(y)) for x in ovoids_p for y in hyp_p}
print "Check H(3): "+ str(min(S) > 3)
```

We find that every pair of ovoidal and semi-singular hyperplane of $H(3)$ through a fixed point which is also the center of the semi-singular hyperplane intersect in more than 3 points. Therefore, every point of \mathcal{S} must be at distance at most 1 from \mathcal{S} . Then it follows from Lemma 4.3.9 that \mathcal{S} is finite.

Finally, let \mathcal{S} be isomorphic to $H(2)$ or $H(4)$ and let q be the order of \mathcal{S} . By Lemma 4.3.8, to show that \mathcal{S} cannot be embedded in any generalized hexagon as a proper full subgeometry, it suffices to check that for every pair of points $x_1, x_2 \in \mathcal{S}$ at distance 3 from each other and for every pair of semi-singular hyperplanes H_1, H_2 with respective centers x_1 and x_2 , we have $|H_1 \cap H_2| > q + 1$. This can be checked using the following SageMath code. Note that by distance transitivity of the automorphism group we only need to check this for one pair of points at distance 3 from each other, thus reducing the amount of computations.


```

# H(2)
points, lines, D = get_data("H2")
p = points[0]
hyp_p = semi_singular_hyperplanes(points, lines, D, p)
dist3 = [x for x in points if D[x][p] == 3]
q = dist3[0]
hyp_q = semi_singular_hyperplanes(points, lines, D, q)
S = {len(set(x) & set(y)) for x in hyp_p for y in hyp_q}
print "Check H(2): " + str(min(S) > 3)

```

```

# H(4)
points, lines D = get_data("H4")
p = points[0]
hyp_p = semi_singular_hyperplanes(points, lines, D, p)
dist3 = [x for x in points if D[x][p] == 3]
q = dist3[0]
hyp_q = semi_singular_hyperplanes(points, lines, D, q)
S = {len(set(x) & set(y)) for x in hyp_p for y in hyp_q}
print "Check H(4): " + str(min(S) > 5)

```

Remark. Generalized hexagons of order greater than 4 seem to be out of reach with our computational methods. And we do not know of any results on intersection sizes of semi-singular and ovoidal hyperplanes of split Cayley hexagons that can help us obtain the above results in general. It would be nice to be able to prove that for all prime powers $q = p^r$, with $p \neq 3$ prime, every pair of semi-singular hyperplanes in $H(q)$ whose centers are at maximum distance 3 intersect in more than $q + 1$ points, which will then imply that these generalized hexagons cannot be contained in bigger generalized hexagons as full subgeometries.

5. Two New Near Octagons

5.1. Introduction

The well-known Hall-Janko near octagon HJ was constructed by Cohen in [52] as the point-line geometry formed by taking a particular conjugacy class C of 315 involutions in the Hall-Janko group J_2 as points, and the sets $\{x, y, xy\}$ where $x, y \in C$ are distinct commuting involutions as lines. This is a regular near octagon of order $(2, 4; 0, 3)$ (see Section 1.3) and has automorphism group isomorphic to $J_2:2$ (the split extension of J_2 by the cyclic group C_2) acting primitively and distance-transitively on the points. A detailed study of this geometry was done by Yoshiara in [142] where in particular it was shown that HJ contains $\text{H}(2)^D$ (and $\text{GO}(2, 1)$) as convex subgeometries. Thus it is natural to wonder whether HJ is the unique near octagon of order $(2, 4)$ which contains $\text{H}(2)^D$ as a full isometrically embedded subgeometry. We address this in Chapter 6, where we give a proof of this fact using the valuation geometry of $\text{H}(2)^D$. In this chapter, we will see how valuation geometries can sometimes give rise to new “interesting” near polygons.

The valuation geometry of HJ is described in Tables 2.7 and 2.8. It is easy to see from the tables that if we take valuations of type A , B and C as points, and all valuation lines of type AAA , ABB , ACC , BBC and CCC as lines (so we skip the valuation lines of type BBB), then we get a partial linear space of order $(2, 10)$ on 4095 points. Using a computer model of this partial linear space, it can be easily checked that this is in fact a *near octagon*. Moreover, computations in SageMath show that the automorphism group G of this near octagon has the following properties:

- G acts transitively on the points;
- $|G| = 503193600$;
- the derived subgroup G' is a simple group of size 251596800.

These properties suggest that G is isomorphic to the group $G_2(4):2$ (see [54]), which is also the automorphism group of the split Cayley hexagon $\text{H}(4)$. We give an alternate construction of this near octagon by directly using the group $G_2(4):2$, which is similar to the construction of HJ given above. This construction will allow us to give a computer-free proof of the fact that the automorphism group of this near octagon is indeed isomorphic to $G_2(4):2$, and help us obtain several useful properties. Recall that a *line spread* of a partial linear space is a set of lines which covers each point exactly once.

Theorem 5.1.1. (1) *Let O_1 be the point-line geometry whose points are the 4095 central involutions¹ of the group $G = G_2(4):2$ and whose lines are all the three element subsets $\{x, y, xy\}$ where x, y are two commuting central involutions that satisfy $[G :$*

¹An involution of a group is called central if its centralizer contains a Sylow 2-subgroup.

$N_G(\langle x, y \rangle) \in \{1365, 13650\}$, with incidence being containment. Then \mathcal{O}_1 is a near octagon of order $(2, 10)$.

- (2) Let S_1 denote the set of all lines $\{x, y, xy\}$ where x and y are two commuting central involutions that satisfy $[G : N_G(\langle x, y \rangle)] = 1365$. Then S_1 is a line spread of \mathcal{O}_1 . If \mathcal{Q}_1 denotes the set of all quads of \mathcal{O}_1 , then for every $Q \in \mathcal{Q}_1$ the lines of S_1 contained in Q define a line spread of Q . Moreover, the point-line geometry \mathcal{S}_1 with point set S_1 , line set \mathcal{Q}_1 and incidence as set containment, is a generalized hexagon isomorphic to the dual split Cayley hexagon $H(4)^D$.

Since $H(4)^D$ contains the generalized hexagon $H(4, 1)$ (which is the unique generalized hexagon of this order) as a subgeometry, we can take a subset S'_1 of the line spread S_1 in Theorem 5.1.1 corresponding to the points of $H(4, 1)$, a subset \mathcal{Q}'_1 of \mathcal{Q}_1 corresponding to the lines of $H(4, 1)$ and then look at the points and lines of the near octagon \mathcal{O}_1 which are contained in elements of S'_1 and \mathcal{Q}'_1 . It turns out that this gives us a subgeometry of \mathcal{O}_1 which is a near octagon of order $(2, 4)$. This near octagon can alternately be described in a similar way as \mathcal{O}_1 .

Theorem 5.1.2. (1) Let \mathcal{O}_2 be the point-line geometry whose points are the 315 central involutions of the group $G = L_3(4):2^{22}$ and whose lines are all the three element subsets $\{x, y, xy\}$ where x, y are two commuting central involutions x, y that satisfy $[G : N_G(\langle x, y \rangle)] \in \{105, 420\}$, with incidence being containment. Then \mathcal{O}_2 is a near octagon of order $(2, 4)$.

- (2) Let S_2 denote the set of all lines $\{x, y, xy\}$ where x and y are two commuting central involutions that satisfy $[G : N_G(\langle x, y \rangle)] = 105$. Then S_2 is a line spread of \mathcal{O}_2 . If \mathcal{Q}_2 denotes the set of all quads of \mathcal{O}_2 , then for every $Q \in \mathcal{Q}_2$ the lines of S_2 contained in Q define a line spread of Q . Moreover, the point-line geometry \mathcal{S}_2 with point set S_2 , line set \mathcal{Q}_2 and incidence as set containment, is isomorphic to the unique generalized hexagon $H(4, 1)$ of order $(4, 1)$.

The near octagons \mathcal{O}_1 and \mathcal{O}_2 will be called the “ $G_2(4)$ near octagon” and the “ $L_3(4)$ near octagon”, respectively. We will prove Theorems 5.1.1(1) and 5.1.2(1) in Section 5.2. The structure of these near octagons around a fixed point can be described by a diagram (see Figures 5.1 and 5.2). These two diagrams are very similar and in Section 5.3 we study geometric properties of a family of near octagons whose local structures can be described by such diagrams. Part (2) of Theorems 5.1.1 and 5.1.2 will follow from that discussion. At this moment, we do not know whether the $G_2(4)$ and $L_3(4)$ near octagons are the only *nontrivial* members of this family of near octagons. In Section 5.4, we study the suboctagons of \mathcal{O}_1 which are isomorphic to HJ , proving that there 416 such suboctagons of \mathcal{O}_1 . We also prove some properties of these suboctagons which will be useful later in Section 5.5 and Chapter 6. Moreover, using these suboctagons we give a new proof of the result of De Wispelaere and Van Maldeghem [70], that HJ^D as a full embedding in the split Cayley hexagon $H(4)$. In Section 5.5 we determine the full automorphism group of the near octagons \mathcal{O}_1 and \mathcal{O}_2 .

Theorem 5.1.3. The automorphism groups of the $G_2(4)$ near octagon and the $L_3(4)$ near octagon are isomorphic to $G_2(4):2$ and $L_3(4):2^2$, respectively.

²It is also standard to denote the group $L_3(4)$ by $PSL_3(4)$, but we will follow the convention of referring to it as $L_3(4)$.

Finally, in Section 5.6 we describe the connection between these near octagons and the distance regular graphs discovered by Soicher in [129], which also gives alternate constructions for these geometries.

Remark. In general, it is hard to construct new “nice near polygons”, for example point-transitive near polygons which are not bipartite graphs. Besides some infinite families (the most recent one being discovered around 15 years ago [56]), there are some examples related to sporadic simple groups which were discovered around 1980 [7], [52], [126]. We hope that our two near polygons may lead to the discovery of some more new “nice near polygons” (for example, by similar constructions using some other group actions).

5.2. Suborbit Diagrams

5.2.1. The $G_2(4)$ Near Octagon

The group $G = G_2(4):2$ has precisely three conjugacy classes of involutions (see [54]). The class $2A$ consists of 4095 involutions all of which are central and contained in the derived subgroup $G' \cong G_2(4)$. A computer model of the group G can be easily constructed using the computer algebra system GAP. All group theoretical claims of the present section have been verified using such a computer model.³

Let \mathcal{P} denote the set of all 4095 central involutions of G . The group G acts on \mathcal{P} by conjugation ($x \mapsto x^g = g^{-1}xg$). Let ω denote a fixed central involution of G . The stabilizer G_ω of ω has eight orbits on \mathcal{P} , the so-called *suborbits of G with respect to ω* . Such a suborbit will be denoted by $\mathcal{O}_0, \mathcal{O}_{1a}, \mathcal{O}_{1b}, \mathcal{O}_{2a}, \mathcal{O}_{2b}, \mathcal{O}_{3a}, \mathcal{O}_{3b}$ and \mathcal{O}_4 in accordance with the information provided by Table 5.1. In the table, we have mentioned the sizes of the suborbits and descriptions for the groups $\langle x, \omega \rangle$, where x is an arbitrary element of the considered suborbit. The suborbits with respect to a certain central involution x will be denoted by $\mathcal{O}_0(x), \mathcal{O}_{1a}(x), \dots, \mathcal{O}_4(x)$.

Suborbit	\mathcal{O}_0	\mathcal{O}_{1a}	\mathcal{O}_{1b}	\mathcal{O}_{2a}	\mathcal{O}_{2b}	\mathcal{O}_{3a}	\mathcal{O}_{3b}	\mathcal{O}_4
Size	1	2	20	40	320	640	1024	2048
$\langle x, \omega \rangle$	C_2	$C_2 \times C_2$	$C_2 \times C_2$	$C_2 \times C_2$	D_8	D_8	S_3	D_{10}

Table 5.1.: Suborbit description for $G_2(4):2$

The central involutions distinct from ω which commute with ω are those of the set $\mathcal{O}_{1a} \cup \mathcal{O}_{1b} \cup \mathcal{O}_{2a}$. Moreover, if $\mathcal{O} \in \{\mathcal{O}_{1a}, \mathcal{O}_{1b}, \mathcal{O}_{2a}\}$ and $x \in \mathcal{O}$, then $\omega x \in \mathcal{O}$. The sets \mathcal{O}_{1a} , \mathcal{O}_{1b} and \mathcal{O}_{2a} consist of those central involutions $x \neq \omega$ for which $[G : N_G(\langle x, \omega \rangle)]$ has size 1365, 13650 and 27300, respectively.

Let \mathcal{L} denote the set of all triples $\{x, y, xy\}$, where x and y are two distinct commuting central involutions such that $[G : N_G(\langle x, y \rangle)]$ has size 1365 or 13650 and let \mathcal{O}_1 be the point-line geometry with point set \mathcal{P} and line set \mathcal{L} , where incidence is containment. Using

³In particular, we do not assume any background on Chevalley groups, as we simply obtain all the required properties of $G_2(4):2$ using elementary arguments along with computations in GAP, and the information given in ATLAS [54].

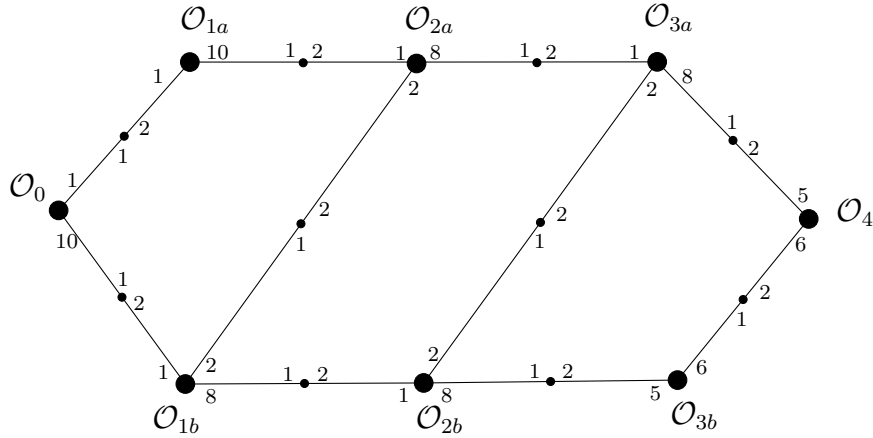


Figure 5.1.: The suborbit diagram for the central involutions of $G_2(4):2$

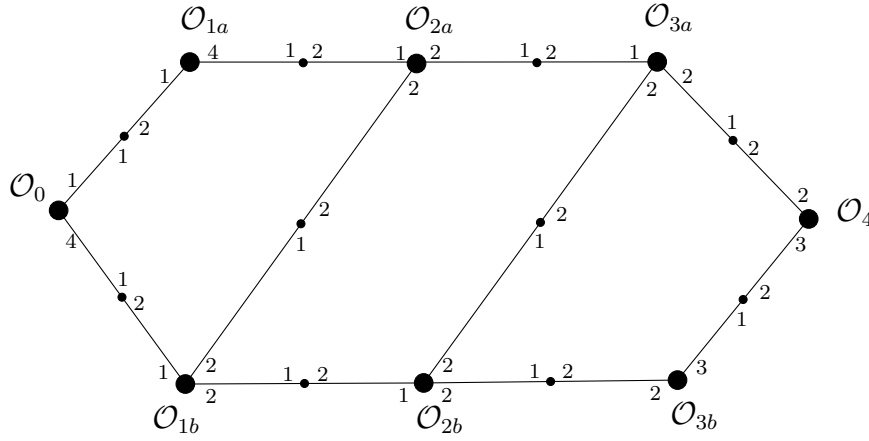
GAP we have computed the suborbit diagram for the central involutions of $G_2(4):2$, which is given in Figure 5.1. Each of the eight big nodes in Figure 5.1 denotes a suborbit and an edge between two such nodes denotes the property that there is a line that intersects both suborbits. A smaller node on each edge denotes a line and the two accompanying numbers denote the number of points of the line that lie in the suborbits it intersects. Each number on a big node denotes the number of lines through a given point in that suborbit going to another suborbit.

In the literature, suborbit diagrams for finite simple groups where adjacency (in the collinearity graph of the involution geometry) is defined by commutativity have been studied (see for example [16], [17] and [142]). For drawing our suborbit diagram(s) we have used similar conventions as in [16]. In our case adjacency involves both commutativity and a condition on the index of certain normalizers. As per [34, Page 397], “Sometimes pairs of commuting involutions from D (where D is a conjugacy class of involutions) fall into several G -orbitals⁴; choosing one orbital for adjacency, one obtains commuting involutions graph in the wider sense”. Therefore, the collinearity graph of the point line geometry \mathcal{O}_1 we have just constructed is a commuting involution graph in the wider sense; noting that we have chosen two suborbits (orbitals) instead of one to define adjacency.

Theorem 5.2.1. *The point-line geometry \mathcal{O}_1 is a near octagon of order $(2, 10)$.*

Proof. Let x be a fixed central involution of $G = G_2(4):2$, i.e., a point of \mathcal{O}_1 . It is clear from the suborbit diagram that every other involution is at distance at most 4 from x . Therefore the point-line geometry is connected and has diameter 4. Now let L be any line, then from the suborbit diagram there exists an $i \in \{0, 1, 2, 3\}$ such that L intersects \mathcal{O}_i in one point and \mathcal{O}_{i+1} in two points. Therefore there exists a unique point on L nearest to x . Since the automorphism group acts transitively on points, \mathcal{O}_1 is a near octagon. Since there are exactly $2 + 20 = 22$ points collinear with x (the elements of $\mathcal{O}_{1a} \cup \mathcal{O}_{1b}$), this near octagon has order $(2, 10)$. \square

⁴orbitals is synonymous with suborbits

Figure 5.2.: The suborbit diagram for the central involutions of $L_3(4):2^2$

5.2.2. The $L_3(4)$ Near Octagon

If instead of the group $G_2(4):2$, we start with $G = L_3(4):2^2$ and take the conjugacy class $2A$ [54] of 315 central involutions of G , then the same construction as before by fixing an element ω in this conjugacy class gives us the suborbit diagram of Figure 5.2. The suborbit description for this case is given in Table 5.2. By a similar argument as before we see that this geometry is a near octagon of order $(2, 4)$, thus proving Theorem 5.1.2(1). We also note that the central involutions distinct from ω which commute with ω are those of the set $\mathcal{O}_{1a} \cup \mathcal{O}_{1b} \cup \mathcal{O}_{2a}$. To gain a better understanding of this $L_3(4)$ near octagon, and to determine its automorphism group, we now describe some of its properties using the projective plane $PG(2, 4)$.

Suborbit	\mathcal{O}_0	\mathcal{O}_{1a}	\mathcal{O}_{1b}	\mathcal{O}_{2a}	\mathcal{O}_{2b}	\mathcal{O}_{3a}	\mathcal{O}_{3b}	\mathcal{O}_4
Size	1	2	8	16	32	64	64	128
$\langle x, \omega \rangle$	C_2	$C_2 \times C_2$	$C_2 \times C_2$	$C_2 \times C_2$	D_8	D_8	S_3	D_{10}

Table 5.2.: Suborbit description for $L_3(4):2$

Let V be a 3-dimensional vector space over the field \mathbb{F}_4 with basis $(\bar{e}_1, \bar{e}_2, \bar{e}_3)$, and denote by $(\bar{f}_1, \bar{f}_2, \bar{f}_3)$ the dual basis in the dual space V' obtained by taking the functional \bar{f}_i 's satisfying the property that $\bar{f}_i(\bar{e}_j) = \delta_{ij}$ for $i, j \in \{1, 2, 3\}$. We get a projective space $PG(V) \cong PG(2, 4)$ from this vector space V by taking the 1-dimensional subspaces of V as points \mathcal{P} , and the 2-dimensional subspaces as lines \mathcal{L} ; incidence is containment. For a 1-dimensional subspace $\langle f \rangle$ of V' , we get a 2-dimensional subspace of V given by $\{u \in V \mid f(u) = 0\}$, implying that we can identify each line of $PG(V)$ with a point of $PG(V')$. A *collineation* of $PG(V) = (\mathcal{P}, \mathcal{L}, I)$ is a bijective map $\theta : \mathcal{P} \cup \mathcal{L} \rightarrow \mathcal{P} \cup \mathcal{L}$ such that $\theta(\mathcal{P}) = \mathcal{P}$, $\theta(\mathcal{L}) = \mathcal{L}$ and $(p, L) \in I$ iff $(p^\theta, L^\theta) \in I$ (incidence preserving). A *correlation* is a bijective incidence preserving map which switches the sets \mathcal{P} and \mathcal{L} .

It is well-known that the group of all collineations of $PG(V)$ is equal to the group $P\Gamma L(V)$. Thus with each collineation θ of $PG(V)$, there is associated a nonsingular 3×3 matrix A over \mathbb{F}_4 and an automorphism τ of \mathbb{F}_4 such that the point $\langle x_1\bar{e}_1 + x_2\bar{e}_2 + x_3\bar{e}_3 \rangle$ of $PG(V)$

is mapped to the point $\langle x'_1 \bar{e}_1 + x'_2 \bar{e}_2 + x'_3 \bar{e}_3 \rangle$ of $\text{PG}(V)$, where $[x'_1 \ x'_2 \ x'_3]^T = A \cdot [x_1^\tau \ x_2^\tau \ x_3^\tau]^T$. While the automorphism τ of \mathbb{F}_4 is uniquely determined by θ , the matrix A itself is only determined up to a nonzero scalar factor. However, all matrices A corresponding to θ have the same determinant as $k^3 = 1$ for every $k \in \mathbb{F}_4 \setminus \{0\}$. The element θ of $\text{P}\Gamma\text{L}(V)$ also permutes the lines of $\text{PG}(V)$. Specifically, if the “line” $\langle y_1 \bar{f}_1 + y_2 \bar{f}_2 + y_3 \bar{f}_3 \rangle$ is mapped to the “line” $\langle y'_1 \bar{f}_1 + y'_2 \bar{f}_2 + y'_3 \bar{f}_3 \rangle$, then $[y'_1 \ y'_2 \ y'_3]^T = (A^T)^{-1} \cdot [y_1^\tau \ y_2^\tau \ y_3^\tau]^T$.

With each correlation θ of $\text{PG}(V)$, there is also associated a nonsingular 3×3 matrix A over \mathbb{F}_4 and an automorphism τ of \mathbb{F}_4 such that the point $\langle x_1 \bar{e}_1 + x_2 \bar{e}_2 + x_3 \bar{e}_3 \rangle$ is mapped to the “line” $\langle y'_1 \bar{f}_1 + y'_2 \bar{f}_2 + y'_3 \bar{f}_3 \rangle$, where $[y'_1 \ y'_2 \ y'_3]^T = A \cdot [x_1^\tau \ x_2^\tau \ x_3^\tau]^T$. Again, the automorphism τ of \mathbb{F}_4 is uniquely determined by θ , but A is only determined up to a nonzero scalar factor. If A and τ are as above, then θ will map the “line” $\langle y_1 \bar{f}_1 + y_2 \bar{f}_2 + y_3 \bar{f}_3 \rangle$ to the point $\langle x'_1 \bar{e}_1 + x'_2 \bar{e}_2 + x'_3 \bar{e}_3 \rangle$, where $[x'_1 \ x'_2 \ x'_3]^T = (A^T)^{-1} \cdot [y_1^\tau \ y_2^\tau \ y_3^\tau]^T$.

Let $G_1 \cong \text{PSL}_3(4) = \text{L}_3(4)$ denote the group of all collineations of $\text{PG}(V)$ for which the field automorphism τ is the identity and the matrix A has determinant 1⁵, let G_2 denote the group of *all* collineations of $\text{PG}(V)$ for which $\det(A) = 1$ and let G denote the group of all collineations and correlations of $\text{PG}(V)$ for which $\det(A) = 1$. Then G_1 has index 2 in G_2 which itself has index 2 in G . If G_3 denotes the set of all elements of G for which $A = I_3$, then $G_3 \cong C_2 \times C_2$. Moreover, G is the internal semidirect product $G_1 \rtimes G_3$. So, G is a group of type $\text{L}_3(4):2^2$. Note also that G_1 is the derived subgroup of G , and that the group of all collineations and correlations of $\text{PG}(V)$ has type $\text{L}_3(4):\text{D}_{12}$.

The little projective group $\text{PSL}_3(4) = \text{L}_3(4)$ is generated by the *elations* of $\text{PG}(2, 4)$, which are collineations that fix every point on a line L , called the *axis* of the elation, and fix every line through a point p on L , called the *center* of the elation. We denote the set of all elations of $\text{PG}(V)$ by Σ . If we treat the axis L as the line at infinity and p as the direction, then the elation $\sigma \in \Sigma$ corresponds to a translation of the affine plane $\text{AG}(2, 4)$ that we get by “removing” the line L . We note the following easily proved properties (where group elements are composed from left to right).

- Lemma 5.2.2.** (1) *If $\sigma \in \Sigma$ is an elation of $\text{PG}(2, 4)$ with center p and axis L , and θ is a collineation of $\text{PG}(2, 4)$, then $\theta^{-1}\sigma\theta$ is an elation of $\text{PG}(2, 4)$ with center p^θ and axis L^θ .*
- (2) *If $\sigma \in \Sigma$ is an elation of $\text{PG}(2, 4)$ with center p and axis L , and θ is a correlation of $\text{PG}(2, 4)$, then $\theta^{-1}\sigma\theta$ is an elation of $\text{PG}(2, 4)$ with center L^θ and axis p^θ .*

Since the characteristic of the field is 2, we see that the (non-identity) elations Σ of $\text{PG}(2, 4)$ form a conjugacy class of involutions in the group generated by collineations and correlations of $\text{PG}(2, 4)$. The elations are all contained in the group $\text{L}_3(4)$, and moreover their conjugacy class is precisely the involution class $2A$ of the group $\text{L}_3(4):2^2$ which has $\text{L}_3(4)$ as its derived subgroup. Therefore, in our construction of the near octagon O_1 , we are taking the 315 elations of the projective plane $\text{PG}(2, 4)$ as points on which the group $G = \text{L}_3(4):2^2$ acts via conjugation. In Section 5.5 we will show that the group G is in fact the full automorphism group of the near octagon O_2 .

We note that the group G has a natural action on the set \mathcal{F} of 105 flags of $\text{PG}(2, 4)$. This action is transitive and there are four suborbits (orbits with respect to the stabilizer

⁵This is well-defined. There are different choices possible for A , but the determinants are always the same.

of a fixed flag), implying that the group G acts distance-transitively on the point set of the unique generalized hexagon $H(4, 1)$ of order $(4, 1)$. Recall that this is the generalized hexagon whose points are the elements of \mathcal{F} and whose lines are the points and lines of $\text{PG}(2, 4)$, with incidence being reverse containment. The fact that G acts distance-transitively implies that G must also act primitively on the set of 105 points of $GH(4, 1)$ (a property that holds for all generalized polygons). Consulting GAP's library of primitive permutation groups [82], we see that there exists a unique primitive permutation group on 105 letters that is of the type $\text{L}_3(4):2^2$. To verify our claims and to construct the suborbit diagram in Figure 5.2, we have used this implementation of the group in GAP.

We now give a computer aided proof of the fact the near octagon \mathcal{O}_2 is a full subgeometry of the near octagon \mathcal{O}_1 .

Theorem 5.2.3. *There exists a full embedding of \mathcal{O}_2 into \mathcal{O}_1 mapping lines of the line spread S_2 to lines of the line spread S_1 . Two embedded points are collinear in \mathcal{O}_1 if and only if they are collinear in \mathcal{O}_2 .*

Proof. By the Atlas [54], the group $G := \text{G}_2(4):2$ has a maximal subgroup H of type $\text{SL}_3(4):2^2$. If we regard G as naturally acting on the dual split Cayley hexagon $H(4)^D$, then H is the stabilizer of a subhexagon of order $(4, 1)$. And in fact this is the presentation of this group and its subgroup which we have used in GAP to verify the claims that follow. Let $Z := Z(H')$ be the center of the derived subgroup $H' \cong \text{SL}_3(4)$ of H . Then $Z \cong C_3$, $J := H/Z \cong \text{L}_3(4):2^2$ and $J' = H'/Z \cong \text{L}_3(4)$. For $x \in H \cong \text{SL}_3(4):2^2$, we let \bar{x} denote the element xZ of $H/Z \cong \text{L}_3(4):2^2$.

There exists a bijective correspondence between the involutions of the group $H' \cong \text{SL}_3(4)$ and the involutions of the group $J' = H'/Z \cong \text{L}_3(4)$. If σ is an involution of H' , then $\sigma \notin Z \cong C_3$ and so $\bar{\sigma}$ is an involution of $J' = H'/Z$. Conversely, if τZ is an involution of $J' = H'/Z$, then there exists a unique involution σ of H' for which $\bar{\sigma} = \tau Z$. Now let Σ denote the set of all involutions of H' for which $\bar{\Sigma} = \{\bar{\sigma} \mid \sigma \in \Sigma\}$ is the conjugacy class of central involutions of J' . Then Σ is a conjugacy class of involutions in both H' and H . Indeed, if $\sigma \in \Sigma$ and $x \in H$, then $\overline{x^{-1}\sigma x} = \bar{x}^{-1} \cdot \bar{\sigma} \cdot \bar{x} \in \bar{\Sigma}$ as $\bar{x} \in H/Z$ and $\bar{\sigma} \in \bar{\Sigma}$, and therefore the involution $x^{-1}\sigma x$ also belongs to Σ .

As Σ is a conjugacy class of involutions in H , it is contained in a conjugacy class of involutions of $G = \text{G}_2(4):2^2$. We have verified with GAP that all elements of Σ are in fact central involutions of G . Since $\overline{x^{-1}\sigma x} = \bar{x}^{-1} \cdot \bar{\sigma} \cdot \bar{x}$ for all $\sigma \in \Sigma$ and all $x \in H$, the action of $H \cong \text{SL}_3(4):2^2$ on the set Σ gives rise to the same suborbit diagram as the action of $H/Z \cong \text{L}_3(4):2^2$ on $\bar{\Sigma}$ (depicted in Figure 5.2). We suppose that \mathcal{O}_2 is the near octagon of order $(2, 4)$ defined on the involutions of the set Σ , where two distinct involutions σ_1 and σ_2 of Σ are collinear whenever the suborbit with respect to σ_1 which contains σ_2 has size 2 or 8. We suppose that \mathcal{O}_1 is the near octagon of order $(2, 10)$ defined on the central involutions of $G = \text{G}_2(4):2$. Since all elements of Σ are central involutions of $\text{G}_2(4):2$, we have identified each point of \mathcal{O}_2 with a point of \mathcal{O}_1 .

Now, fix an involution $\omega \in \Sigma$. We denote by $\mathcal{O}_0, \mathcal{O}_{1a}, \mathcal{O}_{1b}, \mathcal{O}_{2a}, \mathcal{O}_{2b}, \mathcal{O}_{3a}, \mathcal{O}_{3b}$ and \mathcal{O}_4 the suborbits with respect to ω under the action of H using the same notation as before. We also know that the action of the stabilizer of $\text{G}_2(4):2$ with respect to ω on its central involutions gives similar suborbits, which we denote by $\mathcal{O}'_{1a}, \mathcal{O}'_{1b}, \mathcal{O}'_{2a}, \mathcal{O}'_{2b}, \mathcal{O}'_{3a}, \mathcal{O}'_{3b}$ and \mathcal{O}'_4 . With the help of GAP, we have also checked that each of the suborbit \mathcal{O} of Σ is contained in \mathcal{O}' . We know that the suborbits \mathcal{O}_{1a} and \mathcal{O}_{1b} contain those points of \mathcal{O}_2 that

are at distance 1 from ω in \mathcal{O}_2 . We also know that the suborbits \mathcal{O}'_{1a} and \mathcal{O}_{1b} contain those points of \mathcal{O}_1 that are distance 1 from ω in \mathcal{O}_1 . So, if x and y are two points in \mathcal{O}_2 , then we have shown that x and y are collinear in \mathcal{O}_2 if and only if they are collinear in \mathcal{O}_1 . Every line of \mathcal{O}_2 , which is a collection of three mutually collinear points, is thus a line of \mathcal{O}_1 .

The unique line of S_2 through ω is $\{\omega\} \cup \mathcal{O}_{1a}$, which must be equal to $\{\omega\} \cup \mathcal{O}'_{1a}$ ($\mathcal{O}_{1a} = \mathcal{O}'_{1a}$ since they have the same cardinality). Therefore, every line of S_2 is a line of S_1 . \square

5.3. A Family of Near Octagons

Suppose $\mathcal{S} = (\mathcal{P}, \mathcal{L}, \mathcal{I})$ is a finite near octagon of order (s, t) , $s \geq 2$, and S is a line spread of \mathcal{S} . For every point x of \mathcal{S} , let L_x denote the unique line of S containing x . We define a number of additional sets of points of \mathcal{S} .

- For every point x of \mathcal{S} , we define $\Gamma'_1(x) := L_x \setminus \{x\}$ and $\Gamma''_1(x) := \Gamma_1(x) \setminus \Gamma'_1(x)$.
- For every point x of \mathcal{S} and every $i \in \{2, 3\}$, we define $\Gamma'_i(x)$ to be the set of points of $\Gamma_i(x)$ that are collinear with a point of $\Gamma'_{i-1}(x)$, and we put $\Gamma''_i(x) := \Gamma_i(x) \setminus \Gamma'_i(x)$.⁶

Throughout this section, we assume that there exists a positive divisor t' of t , with $t' \neq t$ such that the following hold for every point x of \mathcal{S} :

- (P1) Every point of $\Gamma'_2(x)$ is incident with t' lines meeting $\Gamma''_1(x)$.
- (P2) Every point of $\Gamma''_2(x)$ is incident with a unique line meeting $\Gamma''_1(x)$.
- (P3) Every point of $\Gamma'_3(x)$ is incident with t' lines meeting $\Gamma''_2(x)$.
- (P4) Every point of $\Gamma''_3(x)$ is incident with $\frac{t}{t'}$ lines meeting $\Gamma''_2(x)$.

From the suborbit diagrams of \mathcal{O}_1 and \mathcal{O}_2 (see Figures 5.1, 5.2), described in the previous section, we can see that both of these near octagons satisfies these four properties. For \mathcal{O}_1 , we have $s = 2$, $t = 10$ and $t' = 2$, while for \mathcal{O}_2 we have $s = 2$, $t = 4$ and $t' = 2$. When $t' = 1$, it can be shown that any near polygon \mathcal{S} satisfying these properties must be the direct product of a generalized hexagon of order $(s, t - 1)$ with a line of size $s + 1$. We call this the *trivial* example. Therefore, all the results we prove in this section for the near octagon \mathcal{S} are true for both \mathcal{O}_1 and \mathcal{O}_2 (and for the trivial example). It will also be clear from our discussion that for any point x , the sets $\Gamma_0(x), \Gamma'_1(x), \Gamma''_1(x), \Gamma'_2(x), \Gamma''_2(x), \Gamma'_3(x), \Gamma''_3(x)$ and $\Gamma_4(x)$ correspond to the suborbits $\mathcal{O}_0(x), \mathcal{O}_{1a}(x), \mathcal{O}_{1b}(x), \mathcal{O}_{2a}(x), \mathcal{O}_{2b}(x), \mathcal{O}_{3a}(x), \mathcal{O}_{3b}(x)$ and $\mathcal{O}_4(x)$, respectively.

In the lemmas below, x denotes some fixed point of \mathcal{S} . A point y of \mathcal{S} is said to be of *type* $i \in \{0, 1, 2, 3, 4\}$ if $y \in \Gamma_i(x)$. A point y of \mathcal{S} is said to be of *type* i' or i'' , for $i \in \{1, 2, 3\}$, if it belongs to $\Gamma'_i(x)$ or $\Gamma''_i(x)$, respectively. A line L of \mathcal{S} is said to have *type* (i, j) for $i, j \in \{0, 1', 1'', 2', 2'', 3', 3'', 4\}$ if it contains a unique point of type i and s points of type j ; for example, the line L_x is of type $(0, 1')$. We will ultimately show that, just like the case of near octagons \mathcal{O}_1 and \mathcal{O}_2 , every line of \mathcal{S} has type $(0, 1')$, $(0, 1'')$, $(1', 2')$, $(1'', 2')$, $(1'', 2'')$, $(2', 3')$, $(2'', 3')$, $(2'', 3'')$, $(3', 4)$ or $(3'', 4)$, while every line of the line spread S has type $(0, 1')$, $(1'', 2')$, $(2'', 3')$ or $(3'', 4)$.

⁶ $\Gamma''_2(x)$ and $\Gamma''_3(x)$ are assumed to be nonempty

Lemma 5.3.1. *The point x is contained in a unique line of type $(0, 1')$ and t lines of type $(0, 1'')$. We have $|\Gamma_0(x)| = 1$, $|\Gamma'_1(x)| = s$ and $|\Gamma''_1(x)| = st$. There is a unique line of type $(0, 1')$, namely L_x , and every line of S meeting $\Gamma_0(x) \cup \Gamma'_1(x)$ coincides with L_x .*

Proof. Follows directly from the definitions. \square

Lemma 5.3.2. *Every point $y \in \Gamma'_2(x)$ is contained in a (necessarily unique) quad together with x . This quad has order (s, t') .*

Proof. By definition of the set $\Gamma'_2(x)$, the line $L_x = \{x\} \cup \Gamma'_1(x)$ contains a (necessarily unique) point collinear with y . By Property (P1), there are precisely $t' \geq 1$ points in $\Gamma''_1(x)$ collinear with y . As the points x and y have $t' + 1 \geq 2$ common neighbors (and $s \geq 2$), they are contained in a unique quad, necessarily of order (s, t') , by Theorem 1.3.8. \square

Lemma 5.3.3. *No point of $\Gamma''_2(x)$ can be contained in a quad together with x .*

Proof. Let $y \in \Gamma''_2(x)$. By Property (P2), there is a unique line through y meeting $\Gamma'_1(x)$, and by definition of the set $\Gamma''_2(x)$, there are no lines through y meeting $\Gamma'_1(x)$. Hence, x and y have a unique common neighbor, implying that they cannot be contained in a quad. \square

Lemma 5.3.4. *Every quad through a point y of S contains the line L_y .*

Proof. Without loss of generality, we can take $y = x$. Suppose Q is a quad through x not containing the line L_x . Then no $y \in \Gamma_2(x) \cap Q$ is collinear with a point of L_x , as otherwise L_x will be contained in Q . This implies that $\Gamma_2(x) \cap Q \subseteq \Gamma''_2(x)$, which contradicts Lemma 5.3.3. \square

Lemma 5.3.5. *For every quad Q of S , the lines of the spread S that are contained in Q form a line spread of Q .*

Proof. If this were not the case, then there would exist a line of S meeting Q in a single point y which would contradict Lemma 5.3.4. \square

Lemma 5.3.6. *Through every point $y \in \Gamma'_1(x)$, there is a unique line of type $(0, 1')$ and t lines of type $(1', 2')$.*

Proof. This follows from the definition of the set $\Gamma'_2(x)$; every line through y which is not equal to the unique line of type $(0, 1')$ (the line L_x) must be of type $(1', 2')$. \square

Lemma 5.3.7. *We have $|\Gamma'_2(x)| = s^2t$.*

Proof. By Lemmas 5.3.1 and 5.3.6, the number of edges between $\Gamma'_1(x)$ and $\Gamma'_2(x)$ in the collinearity graph is equal to $|\Gamma'_1(x)| \cdot t \cdot s = s^2t$. As any point of $\Gamma'_2(x)$ is collinear with a unique point of L_x , from double counting these edges we see that $|\Gamma'_2(x)| = s^2t$. \square

Lemma 5.3.8. *There are precisely $\frac{t}{t'}$ quads through any given point of S .*

Proof. Without loss of generality, we prove this for the point x . Every quad through x has order (s, t') and contains s^2t' points at distance 2 from x . All these s^2t' points belong to $\Gamma'_2(x)$ by Lemma 5.3.3. From Lemmas 5.3.2 and 5.3.7, it then follows that there are precisely $\frac{|\Gamma'_2(x)|}{s^2t'} = \frac{s^2t}{s^2t'} = \frac{t}{t'}$ quads through x . \square

Lemma 5.3.9. *Every line M through a point y of S distinct from L_y is contained in a unique quad.*

Proof. Without loss of generality, we may suppose that $y = x$. As any quad through x contains the line L_x by Lemma 5.3.4, there is at most one quad through M (see Lemma 1.3.10). Since every quad through x has order (s, t') , the $\frac{t}{t'}$ quads through x cover $t' \cdot \frac{t}{t'} = t$ lines through x distinct from L_x . As these are all lines through x distinct from L_x , M must be contained in a unique quad. \square

Lemma 5.3.10. *For every $y \in \Gamma_1''(x)$, the line L_y has type $(1'', 2')$. For every $z \in \Gamma_2'(x)$, the line L_z has type $(1'', 2')$.*

Proof. By Lemma 5.3.9, the line xy is contained in a unique quad Q . By Lemma 5.3.5, L_x and L_y are two disjoint lines contained in Q , and so every point of $L_y \setminus \{y\}$ belongs to $\Gamma_2'(x)$, as it is collinear with a point of $L_x \setminus \{x\}$. Therefore, L_y has type $(1'', 2')$. The st mutually distinct lines L_y with $y \in \Gamma_1''(x)$ cover s^2t points of $\Gamma_2'(x)$. By Lemma 5.3.7, these are all the points of $\Gamma_2'(x)$. So, for every $z \in \Gamma_2'(x)$, the line L_z has type $(1'', 2')$. \square

Lemma 5.3.11. *Let $y \in \Gamma_1''(x)$. Then y is contained in a unique line of type $(0, 1'')$, t' lines of type $(1'', 2')$ and $t - t'$ lines of type $(1'', 2'')$.*

Proof. There is a unique line through y containing x , namely the line xy , and this line contains precisely s points of $\Gamma_1''(x)$. There is no line through y meeting $\Gamma_1'(x)$ since otherwise x would be collinear with two distinct points of that line, contradicting the fact that \mathcal{S} is a near polygon. So, every line through y which is distinct from xy meets $\Gamma_2(x)$, and necessarily contains precisely s points of $\Gamma_2(x)$.

By Lemma 5.3.9, there is a unique quad Q (of order (s, t')) through the line xy . By Lemma 5.3.3, the t' lines of Q through y which are distinct from xy all contain precisely s points of $\Gamma_2'(x)$. Conversely, suppose that L is a line through y containing a point $u \in \Gamma_2'(x)$. The unique quad through x and u is a convex subspace and therefore the lines $L = yu$ and xy are contained in this quad, implying that the quad coincides with Q . So, L is one of the t' lines of Q through y distinct from xy . The remaining $t - t'$ lines through y must be of type $(1'', 2'')$. \square

Lemma 5.3.12. *We have $|\Gamma_2''(x)| = s^2t(t - t')$.*

Proof. By Lemmas 5.3.1 and 5.3.11, the number of edges between $\Gamma_1''(x)$ and $\Gamma_2''(x)$ is equal to $|\Gamma_1''(x)| \cdot (t - t')s = s^2t(t - t')$. By Property (P2), we know that the number of edges is also equal to $|\Gamma_2''(x)|$. \square

Lemma 5.3.13. *Let $y \in \Gamma_2'(x)$. Then y is incident with a unique line of type $(1', 2')$, t' lines of type $(1'', 2')$ and $t - t'$ lines of type $(2', 3')$.*

Proof. The lines through y meeting $\Gamma_1(x)$ (necessarily in a unique point) are precisely the $t' + 1$ lines through y that are contained in the unique quad Q through x and y . By Lemma 5.3.3, each of these lines contains precisely s points of $\Gamma_2'(x)$, and hence they have type $(1'', 2')$ or $(1', 2')$. Note that there is a unique line through y meeting $L_x = \{x\} \cup \Gamma_1'(x)$, which is the unique line of type $(1', 2')$ through y . By the definition of the set $\Gamma_3'(x)$, the remaining $t - t'$ lines through y all contain s points of $\Gamma_3'(x)$, and thus have type $(2', 3')$. \square

Lemma 5.3.14. *Through every point $y \in \Gamma_3'(x)$, there is a unique line meeting $\Gamma_2'(x)$.*

Proof. By definition of the set $\Gamma_3'(x)$, we know that there is at least one such line. Suppose two lines through y meet $\Gamma_2'(x)$ in the points u_1 and u_2 , respectively. By Lemma 5.3.13, u_i with $i \in \{1, 2\}$ is collinear with a unique point $v_i \in L_x$. This point necessarily coincides with the unique point v of L_x at distance 2 from y (the near polygon property).

Hence, $v = v_1 = v_2$. Since the points y and v have two distinct neighbors, namely u_1 and u_2 , they are contained in a unique quad Q . The line L_x meets this quad Q in a unique point (namely v), but this contradicts Lemma 5.3.5. \square

Lemma 5.3.15. *We have $|\Gamma'_3(x)| = s^3t(t - t')$.*

Proof. By Lemmas 5.3.7 and 5.3.13, the number of edges between $\Gamma'_2(x)$ and $\Gamma'_3(x)$ is equal to $|\Gamma'_2(x)| \cdot (t - t')s = s^3t(t - t')$. By Lemma 5.3.14, the number of these edges is also equal to $|\Gamma'_3(x)|$. \square

Lemma 5.3.16. *For every $y \in \Gamma''_2(x)$, the line L_y has type $(2'', 3')$. For every $y \in \Gamma'_3(x)$, the line L_y has type $(2'', 3')$.*

Proof. Let $y \in \Gamma''_2(x)$. By Property (P2), there exists a unique point $z \in \Gamma''_1(x)$ collinear with y . By Lemma 5.3.10, the line L_z has type $(1'', 2')$ and so it is distinct from yz . By Lemmas 5.3.4 and 5.3.9, there is a unique quad Q through L_z and yz . By Lemma 5.3.3, this quad cannot contain x . So, x lies at distance 1 from Q and is classical with respect to Q , implying that $L_y \setminus \{y\} \subseteq \Gamma_3(x)$. By Lemma 5.3.5, the line L_y is contained in Q and so the line L_y is parallel and at distance 1 from L_z (see Lemma 1.3.7). Hence, every point of $L_y \setminus \{y\}$ is collinear with a point of $L_z \setminus \{z\} \subseteq \Gamma'_2(x)$, implying that $L_y \setminus \{y\} \subseteq \Gamma'_3(x)$.

The $s^2t(t - t')$ mutually disjoint lines L_z with $z \in \Gamma''_2(x)$ cover $s^3t(t - t')$ points of $\Gamma'_3(x)$. By Lemma 5.3.15, these are all the points of $\Gamma'_3(x)$. So $y \in \Gamma'_3(x)$ and hence the line L_y has type $(2'', 3')$. \square

Lemma 5.3.17. *For any $y \in \Gamma''_3(x) \cup \Gamma_4(x)$, the line L_y has type $(3'', 4)$.*

Proof. This follows from Lemmas 5.3.1, 5.3.10, 5.3.16, and the fact that S is a line spread of \mathcal{S} . \square

Lemma 5.3.18. *There are no lines meeting $\Gamma'_3(x)$ and $\Gamma''_3(x)$.*

Proof. Suppose L is a line containing points $u' \in \Gamma'_3(x)$ and $u'' \in \Gamma''_3(x)$, and let y denote the unique point of $\Gamma_2(x)$ on L . By the definition of $\Gamma''_3(x)$, we must have $y \in \Gamma''_2(x)$ since no point of $\Gamma''_3(x)$ is collinear with a point of $\Gamma'_2(x)$. By Lemma 5.3.14, the point u' is collinear with a unique point $v \in \Gamma'_2(x)$. By Lemma 5.3.10, the line L_v meets $\Gamma''_1(x)$ and so is distinct from vu' . By Lemmas 5.3.4 and 5.3.9, there exists a unique quad Q through L_v and vu' . This quad also contains a point of $\Gamma''_1(x)$, and so the t' lines of Q through u' distinct from vu' all meet $\Gamma_2(x)$. By Lemma 5.3.14, we know that these t' lines meet $\Gamma''_2(x)$. By Property (P3), we know that these are all the lines through u' meeting $\Gamma''_2(x)$. So, the line L is contained in Q . But then every point of $L \setminus \{y\}$ is collinear with a point of $L_v \setminus \Gamma''_1(x) \subseteq \Gamma'_2(x)$, implying that each point of $L \setminus \{y\}$ belongs to $\Gamma'_3(x)$. This contradicts the fact that $u'' \in \Gamma''_3(x)$. \square

The following is a consequence of Lemmas 5.3.1, 5.3.6, 5.3.10, 5.3.11, 5.3.13, 5.3.16, 5.3.17 and 5.3.18.

Corollary 5.3.19. *• Every line of \mathcal{S} has type $(0, 1')$, $(0, 1'')$, $(1', 2')$, $(1'', 2')$, $(1'', 2'')$, $(2', 3')$, $(2'', 3')$, $(2'', 3'')$, $(3', 4)$ or $(3'', 4)$.*

• Every line of the spread S has type $(0, 1')$, $(1'', 2')$, $(2'', 3')$ or $(3'', 4)$.

We can now draw a diagram for \mathcal{S} which looks similar to Figures 5.1 and 5.2. This is done in Figure 5.3, though we still need to prove some of the information that is contained in the diagram.

Lemma 5.3.20. *Let $y \in \Gamma_2''(x)$ and let z be the unique point of $\Gamma_1''(x)$ collinear with y . Then there exists a quad Q through y and z such that the t' lines of Q through y distinct from yz have type $(2'', 3')$.*

Proof. By Property (P2), there is a unique point $z \in \Gamma_1''(x)$ collinear with y . By Lemma 5.3.10, the line L_z contains s points of $\Gamma_2'(x)$. Consider now the unique quad Q through L_z and yz . If L is one of the t' lines of Q through y distinct from yz , then every point of $L \setminus \{y\}$ is collinear with a point of $L_z \setminus \{z\} \subseteq \Gamma_2'(x)$ and hence is contained in $\Gamma_3'(x)$. \square

Lemma 5.3.21. *Every point y of $\Gamma_3'(x)$ is incident with a unique line of type $(2', 3')$, t' lines of type $(2'', 3')$ and $t - t'$ lines of type $(3', 4)$.*

Proof. By Lemma 5.3.14, there is a unique line of type $(2', 3')$ through y and by Property (P3), there are precisely t' lines of type $(2'', 3')$ through y . The remaining $t - t'$ lines through y should all have type $(3', 4)$. \square

Lemma 5.3.22. *Every point y of $\Gamma_2''(x)$ is incident with a unique line of type $(1'', 2'')$, t' lines of type $(2'', 3')$ and $t - t'$ lines of type $(2'', 3'')$.*

Proof. By Lemmas 5.3.12 and 5.3.20, the number of edges between $\Gamma_2''(x)$ and $\Gamma_3'(x)$ is at least $|\Gamma_2''(x)| \cdot st' = s^3 tt'(t - t')$, with equality if and only if every point of type $2''$ is incident with precisely t' lines of type $(2'', 3')$. By Lemma 5.3.15 and Property (P3), we know that the number of edges between $\Gamma_2''(x)$ and $\Gamma_3'(x)$ is precisely $|\Gamma_3'(x)| \cdot t' = s^3 tt'(t - t')$. So, the point $y \in \Gamma_2''(x)$ is incident with precisely t' lines of type $(2'', 3')$. By Property (P2), y is incident with a unique line of type $(1'', 2'')$. The remaining $t - t'$ lines through y should have type $(2'', 3'')$. \square

Lemma 5.3.23. *Every point of $\Gamma_3''(x)$ is contained in $\frac{t}{t'}$ lines of type $(2'', 3'')$ and $t + 1 - \frac{t}{t'}$ lines of type $(3'', 4)$.*

Proof. This follows from Property (P4)⁷ and the discussion so far regarding the types of lines in \mathcal{S} , where we deduced that the only types of lines through a point of type $3''$ are $(2'', 3'')$ and $(3'', 4)$. \square

Lemma 5.3.24. *We have $|\Gamma_3''(x)| = s^3 t'(t - t')^2$.*

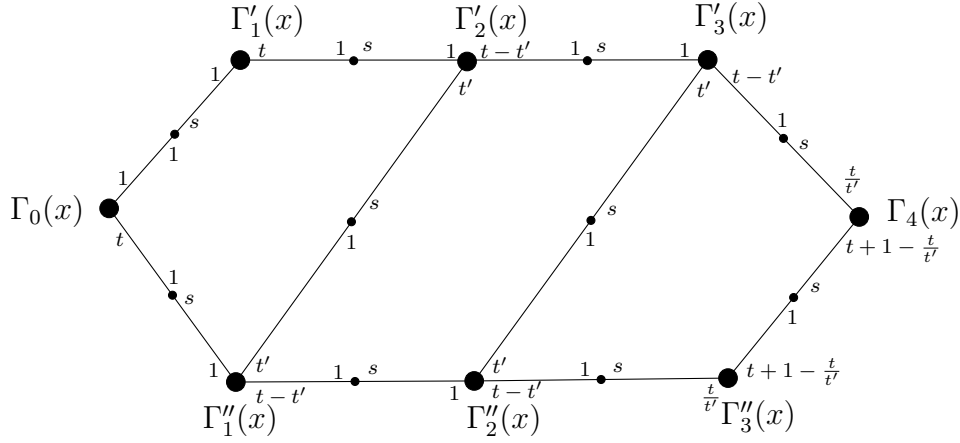
Proof. By Lemmas 5.3.12 and 5.3.22, the number of edges between $\Gamma_2''(x)$ and $\Gamma_3''(x)$ is equal to $|\Gamma_2''(x)| \cdot (t - t')s = s^3 t(t - t')^2$. By Lemma 5.3.23, this number is also equal to $|\Gamma_3''(x)| \cdot \frac{t}{t'}$. Hence, $|\Gamma_3''(x)| = s^3 t'(t - t')^2$. \square

Lemma 5.3.25. *Let L be a line of S meeting $\Gamma_3''(x)$ and $\Gamma_4(x)$. If Q is a quad through L , then Q contains at most one point at distance 2 from x .*

Proof. Since the quad Q contains a point at distance 4 from x , it contains a point at distance 2 from x if and only if x is classical with respect to Q . But then $d(x, Q) = 2$ and so Q contains a unique point at distance 2 from x . \square

Lemma 5.3.26. *Let L be a line of the spread S having type $(3'', 4)$. Then all $\frac{t}{t'}$ quads through L contain a unique point of type $2''$ and a unique line of type $(2'', 3')$. This point of type $2''$ is the unique point of the quad at distance 2 from x , and this line of type $(2'', 3')$ belongs to S .*

⁷This is the first time we have used (P4).

Figure 5.3.: The structure of \mathcal{S} with respect to the fixed point x

Proof. Let y be the unique point of L contained in $\Gamma''_3(x)$. By Property (P4), there are $\frac{t}{t'}$ lines K through y meeting $\Gamma''_2(x)$ giving rise to $\frac{t}{t'}$ distinct quads through L , each of which contains a unique point of $\Gamma''_2(x)$ by Lemma 5.3.25. These are all the quads through L . Let Q be any of these $\frac{t}{t'}$ quads and let z denote the unique point in $Q \cap \Gamma_2(x)$. By Lemmas 5.3.20 and 5.3.22, the t' lines of type $(2'', 3')$ through z are all contained in a quad Q' through z . As Q' contains a point of $\Gamma'_1(x)$, we have $Q \neq Q'$ and so $Q \cap Q'$ is at most a line. As $L_z \subseteq Q \cap Q'$, we have $L_z = Q \cap Q'$. Clearly, $L_z = Q \cap Q'$ is the unique line of type $(2'', 3')$ of Q . \square

Lemma 5.3.27. *Every point y of $\Gamma_4(x)$ is incident with $\frac{t}{t'}$ lines of type $(3', 4)$ and $t+1-\frac{t}{t'}$ lines of type $(3'', 4)$.*

Proof. If L is a line through y meeting $\Gamma'_3(x)$, then Lemmas 5.3.4, 5.3.9 and 5.3.17 imply that there is a unique quad through L and this quad contains the line L_y of type $(3'', 4)$. Conversely, by Lemma 5.3.26, any of the $\frac{t}{t'}$ quads through L_y contains a unique line through y meeting $\Gamma'_3(x)$. So, there are $\frac{t}{t'}$ lines through y meeting $\Gamma'_3(x)$ and $t+1-\frac{t}{t'}$ lines meeting $\Gamma''_3(x)$. \square

Lemma 5.3.28. *We have $|\Gamma_4(x)| = s^4 t' (t - t')^2$.*

Proof. By Lemmas 5.3.15 and 5.3.21, the number of edges between $\Gamma'_3(x)$ and $\Gamma_4(x)$ is equal to $|\Gamma'_3(x)| \cdot (t - t')s = s^4 t (t - t')^2$. By Lemma 5.3.27, the number of such edges is also equal to $|\Gamma_4(x)| \cdot \frac{t}{t'}$. It follows that $|\Gamma_4(x)| = s^4 t' (t - t')^2$. \square

We now have a full description of \mathcal{S} in Figure 5.3, where we follow a similar convention as in Figures 5.1 and 5.2. Note that we have $|\Gamma_0(x)| = 1$, $|\Gamma'_1(x)| = s$, $|\Gamma''_1(x)| = st$, $|\Gamma'_2(x)| = s^2 t$, $|\Gamma''_2(x)| = s^2 t (t - t')$, $|\Gamma'_3(x)| = s^3 t (t - t')$, $|\Gamma''_3(x)| = s^3 t' (t - t')^2$, and $|\Gamma_4(x)| = s^4 t' (t - t')^2$. Therefore, the total number of points in \mathcal{S} can also be given in terms of the parameters s, t, t' .

Lemma 5.3.29. *Every point-quad pair in \mathcal{S} is classical.*

Proof. Without loss of generality, we may suppose that x is the point and Q is a quad of \mathcal{S} . If $d(x, Q) = 1$, then by Lemma 1.3.11 the pair (x, Q) is classical. By Lemma 5.3.26, (x, Q) is classical if Q contains points at distance 4 from x . So, suppose $Q \subseteq \Gamma_2(x) \cup \Gamma_3(x)$. Note that $Q \cap \Gamma_2(x)$ must be nonempty. Let $y \in Q \cap \Gamma_2(x)$. Then $L_y \subseteq Q$ implies by

Lemma 5.3.10 that $y \notin \Gamma'_2(x)$. So, $y \in \Gamma''_2(x)$. The line L_y meets $\Gamma'_3(x)$. By Property (P2), y is collinear with a unique point z of $\Gamma'_1(x)$. Denote by Q' the unique quad through yz and L_y . By Lemmas 5.3.4 and 5.3.9, Q' is the unique quad through y meeting $\Gamma''_1(x)$. By Lemmas 5.3.20 and 5.3.22, the t' lines of type $(2'', 3')$ through y are the t' lines of Q' through y distinct from yz . Each of the $t - t'$ lines through y meeting $\Gamma''_3(x)$ determines a quad together with L_y and such a quad has t' lines through y meeting $\Gamma''_3(x)$. So, there are $\frac{t-t'}{t'} = \frac{t}{t'} - 1$ quads through L_y meeting $\Gamma''_3(x)$ and hence also $\Gamma_4(x)$ as the unique line of S through a point of $\Gamma''_3(x)$ meets $\Gamma_4(x)$. We see that we have accounted for all $\frac{t}{t'}$ quads through y , and none of these quads can be equal to Q , which is contained in $\Gamma_2(x) \cup \Gamma_3(x)$. \square

Lemma 5.3.30. *Two distinct lines L_1 and L_2 of the spread S lie at distance 1 from each other if and only if they are contained in the same quad. If $d(L_1, L_2) = 1$, then L_1 and L_2 are parallel.*

Proof. If L_1 and L_2 are contained in a quad, then they are parallel and $d(L_1, L_2) = 1$. Conversely, suppose that $d(L_1, L_2) = 1$ and let $x_1 \in L_1$ and $x_2 \in L_2$ such that $d(x_1, x_2) = 1$. The unique quad through the line x_1x_2 then contains L_1 and L_2 . \square

Lemma 5.3.31. *Every two lines L and L' of S are parallel.*

Proof. Let $x \in L$ and $x' \in L'$ such that $d(L, L') = d(x, x')$. It suffices to show that every point of L has distance at most $\delta := d(x, x')$ from L' . Let $x = y_0, y_1, \dots, y_\delta = x'$ be a path connecting x and x' . Put $L_i := L_{y_i}$, $i \in \{0, 1, \dots, \delta\}$. Then for every $i \in \{1, 2, \dots, \delta\}$, the lines L_{i-1} and L_i are either equal or parallel at distance 1. So, every point of L_{i-1} has distance at most 1 from a point of L_i . This implies that every point of $L = L_0$ has distance at most δ from $L' = L_\delta$. \square

Let \mathcal{Q} denote the set of quads of \mathcal{S} , and let \mathcal{S}' be the point-line geometry with point set S and line set \mathcal{Q} , where incidence is containment.

Lemma 5.3.32. *If $L_1, L_2 \in S$, then the distance d between L_1 and L_2 in \mathcal{S}' is equal to $d(L_1, L_2)$.*

Proof. Put $d_1 := d$ and $d_2 := d(L_1, L_2)$. Suppose $L_1 = K_0, K_1, \dots, K_{d_1} = L_2$ is a (shortest) path in \mathcal{S}' connecting L_1 and L_2 . Then the K_i 's are mutually disjoint. Let x_0 be an arbitrary point of K_0 and for every $i \in \{1, 2, \dots, d_1\}$, let x_i denote the unique point of K_i collinear with x_{i-1} . Then $x_0 \in L_1$ lies at distance at most d_1 from $x_{d_1} \in L_2$, showing that $d_2 \leq d_1$.

Suppose y_0, y_1, \dots, y_{d_2} is a (shortest) path of length d_2 in \mathcal{S} connecting a point $y_0 \in L_1$ with a point $y_{d_2} \in L_2$. In the sequence $L_{y_0}, L_{y_1}, \dots, L_{y_{d_2}}$, any two consecutive lines are either equal or at distance 1 from each other (in \mathcal{S}') by Lemma 5.3.30. So, we should also have that $d_1 \leq d_2$. We conclude that $d_1 = d_2$. \square

Lemma 5.3.33. *\mathcal{S}' has order $(st', \frac{t}{t'} - 1)$.*

Proof. By Lemma 5.3.5, every line of \mathcal{S}' contains precisely $1 + st'$ points (this the size of a line spread of a generalized quadrangle of order (s, t')). By Lemmas 5.3.4 and 5.3.8, every line of S is contained in precisely $\frac{t}{t'}$ quads, showing that every point of \mathcal{S}' is incident with precisely $\frac{t}{t'}$ lines of \mathcal{S}' . \square

Lemma 5.3.34. *\mathcal{S}' is a near hexagon.*

Proof. Let $L \in S$ and $Q \in \mathcal{Q}$. Denote by S_Q the set of lines of S contained in Q . By Lemma 5.3.5, S_Q is a line spread of the generalized quadrangle Q . Let x be an arbitrary point of L . As x is classical with respect to the quad Q , there exists a unique point $x' \in Q$ nearest to x . By Lemma 5.3.32, $L_{x'}$ is the unique line of S_Q at distance $d(x, x')$ from $L_x = L$ in the geometry \mathcal{S}' . Every other line of S_Q has distance $d(x, x') + 1$ from $L_x = L$.

It remains to show that \mathcal{S}' has diameter 3. By Lemma 5.3.32, the diameter of \mathcal{S}' is at most 3. If x_1 and x_2 are two points of \mathcal{S} at distance 4 from each other, then the fact that L_{x_1} and L_{x_2} are parallel implies that they lie at distance 3 from each other (both in \mathcal{S} as \mathcal{S}'). So, the diameter of \mathcal{S}' is indeed 3. \square

Lemma 5.3.35. *Every two points of \mathcal{S}' at distance 2 from each other have a unique common neighbor.*

Proof. Let $L_1, L_2 \in S$ be at distance 2 from each other in the geometry \mathcal{S}' and suppose $M_1, M_2 \in S$ are two distinct neighbors of L_1, L_2 . The lines L_1, L_2, M_1, M_2 are mutually parallel by Lemma 5.3.31. By Lemma 5.3.32, there exist points $x_1 \in L_1$ and $x_2 \in L_2$ at distance 2 from each other. Also, by Lemma 5.3.32, the points x_1, x_2 have at least two common neighbors, one on M_1 and another one on M_2 . So, x_1 and x_2 are contained in a quad Q . This quad should contain the line $L_{x_1} = L_1$, but that is impossible since $d(x_2, L_1) = 2$. \square

The following is an immediate consequence of Lemmas 5.3.33, 5.3.34 and 5.3.35.

Corollary 5.3.36. *\mathcal{S}' is a generalized hexagon of order $(st', \frac{t}{t'} - 1)$.*

Since there is a unique generalized hexagon of order $(4, 1)$, we have proved Theorem 5.1.2 completely. For Theorem 5.1.1(1), we still need to check if the generalized hexagon \mathcal{S}_1 of order 4 is isomorphic to the dual split Cayley hexagon $H(4)^D$. This can be done easily in a computer model of \mathcal{S}_1 using the geometrical characterization of dual Split Cayley hexagons obtained by Ronan [120].

5.4. Hall-Janko Suboctagons of O_1

In this section, we construct and classify all Hall-Janko suboctagons of O_1 . These are (full) subgeometries of O_1 that are isomorphic to HJ. We will show that there are 416 such subgeometries and that all of them are isometrically embedded. In the following lemma, we already construct all these 416 subgeometries from the 416 (maximal) subgroups of $G_2(4):2$ isomorphic to $J_2:2$. But first we include some group theoretical properties of the group $G_2(4):2$ and $J_2:2$ that one can obtain from [54] and GAP.

The group $G_2(4):2$ has $J_2:2$ as a maximal subgroup of index 416. If we have $H \cong J_2:2$ and $G \cong G_2(4):2$ such that $H < G$ and if Σ_H, Σ_G denote the corresponding conjugacy classes of central involutions, then $\Sigma_H = H \cap \Sigma_G$. Moreover, $\Sigma_H \subseteq H' \cong U_3(3)$ (the derived subgroup), and $\Sigma_G \subseteq G' \cong G_2(4)$. We also note that if x, y are two distinct commuting central involutions of H , then $[G : N_G(\langle x, y \rangle)] = 13650$. Since Σ_G generates a normal subgroup of the simple group G' , we necessarily have $\langle \Sigma_G \rangle = G'$. Similarly, $\langle \Sigma_H \rangle = H'$. Also, $N_G(H) = H$, and hence H has precisely 416 conjugates in G . There exists a natural

bijjective correspondence between the subgroups of $G = G_2(4):2$ isomorphic to $J_2:2$ and the subgroups isomorphic to J_2 . Every subgroup isomorphic to $J_2:2$ contains a unique J_2 -subgroup, namely its derived subgroup. Conversely, every J_2 -subgroup H of G must be contained in a unique (maximal) subgroup isomorphic to $J_2:2$, the normalizer of H inside G .

Lemma 5.4.1. (1) *Let H be a (maximal) subgroup of $G = G_2(4):2$ isomorphic to $J_2:2$. Then the set Σ_H of the central involutions of G contained in H is a subspace of the near octagon \mathcal{O}_1 on which the induced subgeometry, denoted by \mathcal{S}_H , is isomorphic to the Hall-Janko near octagon \mathcal{HJ} .*

(2) *If H_1 and H_2 are two distinct maximal subgroups of G isomorphic to $J_2:2$, then \mathcal{S}_{H_1} and \mathcal{S}_{H_2} are distinct subgeometries.*

Proof. (1) On the set $\Sigma_H \subseteq H' \cong J_2$, a Hall-Janko near octagon \mathcal{S}'_H can be defined by taking as lines all the sets $\{x, y, xy\}$, where x and y are two distinct commuting elements of Σ_H . Recall that if the elements $x, y \in \Sigma_H$ commute, then $[G : N_G(\langle x, y \rangle)] = 13650$, implying that $\{x, y, xy\}$ is a line of \mathcal{O}_1 , with $y \in \mathcal{O}_{1b}(x)$ (see Section 5.2). Conversely, if $x, y \in \Sigma_H$ such that $\{x, y, xy\}$ is a line of \mathcal{O}_1 , then x, y commute and hence $\{x, y, xy\}$ is also a line of \mathcal{S}'_H .

(2) We need to show that H is uniquely determined by Σ_H . The subgroup generated by Σ_H is a normal subgroup of $H' \cong J_2$ and hence coincides with H' . Inside $G = G_2(4) : 2$, there is a unique subgroup isomorphic to $J_2 : 2$ that contains $H' \cong J_2$, namely its normalizer. Hence, $H = N_G(\langle \Sigma_H \rangle)$. □

Before proceeding to prove that every Hall-Janko suboctagon is as described in Lemma 5.4.1, we first give an alternative proof of a result of De Wispelaere and Van Maldeghem [70].

Lemma 5.4.2. *The geometry \mathcal{HJ}^D has a full embedding in $\mathcal{H}(4)$.*

Proof. Let H be a maximal subgroup of $G = G_2(4):2$ isomorphic to $J_2:2$. Then by Lemma 5.4.1, $\mathcal{S}_H \cong \mathcal{HJ}$ is a full subgeometry of \mathcal{O}_1 . Note that no line of \mathcal{S}_H is contained in the line spread S_1 of \mathcal{O}_1 as pairs of collinear points x, y in \mathcal{S}_H satisfy $[G : N_G(\langle x, y \rangle)] = 13650$. Therefore, every line L of \mathcal{S}_H is contained in a unique quad Q_L by Lemma 5.3.9. As any two involutions of $Q_L \cap H$ commute, $Q_L \cap H$ is at most a line of H , implying that $L = Q_L \cap H$. So, if L_1, L_2, \dots, L_5 are the five lines of \mathcal{S}_H through a given point x , then the quads $Q_{L_1}, Q_{L_2}, \dots, Q_{L_5}$ are mutually distinct and hence are all the five quads through the line L_x of the spread S_1 (see Lemmas 5.3.4 and 5.3.8). This implies that the maps $x \mapsto L_x, L \mapsto Q_L$ define a full embedding of the dual of \mathcal{S}_H into the dual of (S_1, \mathcal{Q}_1) , which is isomorphic to $\mathcal{H}(4)$ (see Theorem 5.1.1(2)). □

In the rest of this section, \mathcal{H} will denote an arbitrary Hall-Janko suboctagon, not necessarily arising from the central involutions contained in a Hall-Janko subgroup. We will derive several properties of \mathcal{H} that will enable us to prove that there are at most (and hence precisely) 416 Hall-Janko suboctagons of \mathcal{O}_1 . For $x, y \in \mathcal{O}_1$, we simply denote $d_{\mathcal{O}_1}(x, y)$ by $d(x, y)$.

Lemma 5.4.3. *If x and y are two points of \mathcal{H} such that $d_{\mathcal{H}}(x, y) \leq 2$ then $d_{\mathcal{H}}(x, y) = d(x, y)$.*

Proof. Follows directly from the fact that \mathcal{H} is a full subgeometry of the near polygon \mathcal{O}_1 . \square

Lemma 5.4.4. *If x and y are two points of \mathcal{H} such that $d_{\mathcal{H}}(x, y) = d(x, y) = 3$ then $y \in \mathcal{O}_{3b}(x)$.*

Proof. Since \mathcal{H} is a regular near octagon with parameters $(2, 4; 0, 3)$ there must be four lines of \mathcal{H} through y that contain a point at distance 2 from x in \mathcal{H} (and hence also in \mathcal{O}_1 by Lemma 5.4.3). But, by the suborbit diagram, if y lies in $\mathcal{O}_{3a}(x)$ then there are only three lines through y containing a point at distance 2 from x (see Figure 5.1). \square

Lemma 5.4.5. *Let Q be a quad of \mathcal{O}_1 and x, y two points of Q such that $d(x, y) = 2$. If z is a point collinear with y and not contained in Q then $z \in \mathcal{O}_{3a}(x)$.*

Proof. As z is classical with respect to Q , $d(x, z) = 3$. Now this directly follows from the suborbit diagram with respect to x and Lemma 5.3.3. \square

Lemma 5.4.6. *A quad Q of \mathcal{O}_1 cannot contain a pair of intersecting lines of \mathcal{H} .*

Proof. Suppose L_1 and L_2 are two intersecting lines of \mathcal{H} contained in Q . Let $x_1 \in L_1 \setminus L_2$ and $x_2 \in L_2 \setminus L_1$. As there are five lines through x_2 contained in \mathcal{H} and only three contained in Q , there exists a neighbor x_3 of x_2 in $\mathcal{H} \setminus Q$. For this point x_3 , we have $d_{\mathcal{H}}(x_1, x_3) = d(x_1, x_3) = 3$. By Lemma 5.4.5 $x_3 \in \mathcal{O}_{3a}(x_1)$. This contradicts Lemma 5.4.4 which would imply that $x_3 \in \mathcal{O}_{3b}(x_1)$. \square

Lemma 5.4.7. *None of the lines of the line spread S_1 is contained in \mathcal{H} .*

Proof. Suppose L is a line of S_1 contained in \mathcal{H} , and let M denote any other line of \mathcal{H} meeting L in a point. By Lemmas 5.3.4 and 5.3.9, there is a unique quad Q containing L and M . This quad would contradict Lemma 5.4.6. \square

Lemma 5.4.8. *Every quad which contains a point x of \mathcal{H} contains a unique line of \mathcal{H} through x .*

Proof. By Lemma 5.4.7, the line L_x is not contained in \mathcal{H} . There are five lines through x contained in \mathcal{H} . By Lemma 5.4.6, each of the five quads through L_x contains at most one and hence precisely one of these five lines. \square

Lemma 5.4.9. *\mathcal{H} is isometrically embedded into \mathcal{O}_1 .*

Proof. Suppose $d(x, y) \neq d_{\mathcal{H}}(x, y)$ for certain points x and y of \mathcal{H} , and suppose x and y have been chosen in such a way that $i := d_{\mathcal{H}}(x, y)$ is as small as possible. By Lemma 5.4.3, $i \in \{3, 4\}$. Let y' be a point in \mathcal{H} such that $d_{\mathcal{H}}(x, y') = i - 1$, $d_{\mathcal{H}}(y', y) = 1$ and let y'' denote the third point on the line yy' . By the near polygon property we know that $d_{\mathcal{H}}(x, y'') = d_{\mathcal{H}}(x, y) = i$. By the minimality of i , $d(x, y') = i - 1$. Again by the near polygon property, $\{d(x, y), d(x, y'')\} = \{i - 1, i - 2\}$. So, still under the assumption that the distance $d_{\mathcal{H}}(x, y)$ is as small as possible, we could have chosen y in such a way that $d(x, y) = d_{\mathcal{H}}(x, y) - 2$.

Suppose $i = 3$. Then we have $x, y \in \mathcal{H}$ such that $d_{\mathcal{H}}(x, y) = 3$ and $d(x, y) = 1$. Let x, z_1, z_2, y be a shortest path between x and y in \mathcal{H} . By Lemma 5.4.3, $d(x, z_2) = 2$. Since x and z_2 have at least two common neighbors in \mathcal{G} (namely y and z_1), there exists a quad Q containing x, z_2 and all their common neighbors. The quad Q would then contain the intersecting lines xz_1 and z_1z_2 , which is in contradiction with Lemma 5.4.6. Therefore, if x and y are two points of \mathcal{H} with $d_{\mathcal{H}}(x, y) \leq 3$, then $d_{\mathcal{H}}(x, y) = d(x, y)$.

Now suppose $i = 4$, i.e., we have $x, y \in \mathcal{H}$ such that $d_{\mathcal{H}}(x, y) = 4$ and $d(x, y) = 2$. Let y' be a common neighbor of x and y in \mathcal{O}_1 . There exists a quad Q through the line xy' which by Lemma 5.4.8 must contain a line M of \mathcal{H} through x . Let x' be the unique point on M satisfying $d_{\mathcal{H}}(y, x') = 3$. From the previous case, we also have $d(y, x') = 3$. So, $y \in \mathcal{O}_{3b}(x')$ by Lemma 5.4.4. Since $d(y, x') = 3$, the quad Q cannot contain the point y . Lemma 5.4.5 (with x, y and z replaced by x', y' and y) would then imply that $y \in \mathcal{O}_{3a}(x')$, which is in contradiction with the earlier claim that $y \in \mathcal{O}_{3b}(x')$. \square

Lemma 5.4.10. *If x and y are two points of \mathcal{H} such that $d_{\mathcal{H}}(x, y) = 2$ then $y \in \mathcal{O}_{2b}(x)$.*

Proof. We also have $d(x, y) = 2$. Let $x' \in \mathcal{H}$ be a common neighbor of x and y . If $y \in \mathcal{O}_{2a}(x)$, then the unique quad through x and y would contain the intersecting lines xx' and $x'y$, which would be in violation with Lemma 5.4.6. Therefore, $y \in \mathcal{O}_{2b}(x)$. \square

Lemma 5.4.11. *Through every pair of opposite points of \mathcal{O}_1 there is at most one Hall-Janko suboctagon.*

Proof. Let x and y be two opposite points of \mathcal{O}_1 and suppose a Hall-Janko suboctagon \mathcal{H} contains x and y . We will show that \mathcal{H} is uniquely determined by x and y . In this proof all suborbits are considered with respect to the point x . By Lemma 5.4.9, the distance between two points of \mathcal{H} is the same in the geometries \mathcal{H} and \mathcal{O}_1 .

There are five lines through y inside \mathcal{H} that contain a point at distance 3 from x . By Lemma 5.4.4 all of these lines must intersect \mathcal{O}_{3b} . By the suborbit diagram and Lemma 5.3.17, there are exactly six such lines through y and one of them is in the line spread S_1 . By Lemma 5.4.7, the line belonging to S_1 cannot be contained in \mathcal{H} . Therefore the five lines of \mathcal{H} through y , going back to x are uniquely determined by x and y .

Now let $y' \in \mathcal{O}_{3b}$ be a point on one of these five lines and Q the unique quad through yy' and $L_{y'} \neq yy'$. By Lemma 5.3.17, $L_{y'}$ meets \mathcal{O}_4 . By Lemma 5.4.8 the third line of Q through y' , call it $M_{y'}$, doesn't lie in \mathcal{H} . We claim that $M_{y'}$ intersects \mathcal{O}_{2b} . Indeed, as the point x is classical with respect to Q , the unique point u in Q nearest to x lies at distance 2 from x and is collinear with y' . Therefore, $u \in \mathcal{O}_{2b}$ and $M_{y'} = y'u$. The four lines of \mathcal{H} through y' that go back to x are now uniquely determined. Indeed, by Lemma 5.4.10, each of the four lines of \mathcal{H} through y' meets \mathcal{O}_{2b} . But by the suborbit diagram, there are precisely five such lines. Moreover, one of these five lines is the line $M_{y'}$ and we already know that it cannot be a line of \mathcal{H} .

Now, let $y'' \in \mathcal{O}_{2b}$ be a point on one of these four lines. By the suborbit diagram there is a unique line through y'' containing a point y''' in \mathcal{O}_{1b} , which must necessarily be in \mathcal{H} . Moreover, there is a unique line through y''' that contains x .

So far, we have proved that given any point y in \mathcal{H} with $d_{\mathcal{H}}(x, y) = 4$, all shortest paths between x and y in \mathcal{H} are uniquely determined by x and y . Moreover, all points at distance 4 from x that are collinear with y are uniquely determined. These properties in fact imply that the whole of \mathcal{H} is uniquely determined. Indeed, the subgraph of the collinearity graph induced on the set $\Gamma_4(x) \cap \mathcal{H}$ is connected (see Step 1 of the proof of Theorem 3 in [53]), and every shortest path between x and a point of \mathcal{H} can be extended to a shortest path between x and a point of $\Gamma_4(x) \cap \mathcal{H}$. \square

Lemma 5.4.12. *There are precisely 416 Hall-Janko suboctagons of \mathcal{O}_1 , namely the suboctagons \mathcal{S}_H for maximal subgroups $H \cong J_2:2$ of $G = G_2(4):2$. Through every pair of opposite points of \mathcal{G} , there is precisely one Hall-Janko suboctagon.*

Proof. A Hall-Janko suboctagon has $315 \cdot 64$ ordered pairs of opposite points while \mathcal{O}_1 has $4095 \cdot 2048$ such pairs. Therefore by Lemma 5.4.11, there are at most $(4095 \cdot 2048)/(315 \cdot 64) = 416$ Hall-Janko suboctagons in \mathcal{O}_1 . By Lemma 5.4.1 there are at least that many. \square

Now that we have classified all Hall-Janko sub near octagons of \mathcal{O}_1 , we end this section by proving some extra properties of these Hall-Janko suboctagons.

Lemma 5.4.13. *If \mathcal{H} is a Hall-Janko suboctagon of \mathcal{O}_1 and x a point not contained in \mathcal{H} , then there is a unique point x' in \mathcal{H} that is collinear with x .*

Proof. Let \mathcal{H} be a Hall-Janko suboctagon of \mathcal{O}_1 and x a point not contained in \mathcal{H} . Say x has two neighbors y, z in \mathcal{H} . Then by Lemma 5.4.9 $d_{\mathcal{H}}(y, z) = 2$ and hence there is a common neighbor of y, z inside \mathcal{H} . This means that there is a quad through y, z whose intersection with \mathcal{H} contains a pair of intersecting lines, contradicting Lemma 5.4.6. Therefore, if x has a neighbor in \mathcal{H} then it has a unique neighbor.

Now we can show that x has a neighbor in \mathcal{H} by a simple counting. There are six lines out of the eleven through each point in \mathcal{H} that are not contained in \mathcal{H} , giving us a total of $12 \cdot 315$ points of \mathcal{O}_1 at distance 1 from \mathcal{H} . Adding this to the number of points in \mathcal{H} we get $315 \cdot 12 + 315 = 4095$ which is the total number of points in \mathcal{O}_1 . \square

For a Hall-Janko suboctagon \mathcal{H} and a point x of \mathcal{O}_1 we define the projection of x onto \mathcal{H} , $\pi_{\mathcal{H}}(x)$, to be x if $x \in \mathcal{H}$ and the unique point $x' \in \mathcal{H}$ collinear with x if $x \notin \mathcal{H}$.

Lemma 5.4.14. *Let \mathcal{H} be a Hall-Janko suboctagon of \mathcal{O}_1 and x, y be two distinct points not contained in \mathcal{H} such that $\pi_{\mathcal{H}}(x) = \pi_{\mathcal{H}}(y)$. Then $\mathcal{H} \cap \Gamma_4(x) \neq \mathcal{H} \cap \Gamma_4(y)$.*

Proof. We consider the following two cases:

- (1) The point x is collinear with y . Let $x' = \pi_{\mathcal{H}}(x) = \pi_{\mathcal{H}}(y)$ and $z \in \Gamma_4(x') \cap \mathcal{H}$. Since $\{x, y, x'\}$ is a line, either $d(z, x) = 3$ and $d(z, y) = 4$, or $d(z, x) = 4$ and $d(z, y) = 3$. In either case z belongs to only one of $\mathcal{H} \cap \Gamma_4(x)$, $\mathcal{H} \cap \Gamma_4(y)$.
- (2) The point x is not collinear with y . Consider the suborbit diagram with ω equal to the common projection of x and y .

Let z be a point in $\mathcal{O}_{3b} \cap \mathcal{H}$. There are five lines through z going back to \mathcal{O}_{2b} and four of them are contained in \mathcal{H} . The one line that is not contained in \mathcal{H} gives us a unique point z' of $\mathcal{O}_{2b} \setminus \mathcal{H}$ collinear with z . This in turn gives us a unique point u in $\mathcal{O}_{1b} \setminus \mathcal{H}$ collinear with ω and z' . This point has distance 2 from z and cannot belong to \mathcal{H} by Lemma 5.4.13.

Conversely, let u be a point in $\mathcal{O}_{1b} \setminus \mathcal{H}$. It has sixteen neighbors in \mathcal{O}_{2b} none of which is contained in \mathcal{H} by Lemma 5.4.13. By the suborbit diagram and Lemma 5.4.4, the projection of each of these sixteen points in \mathcal{H} must lie in \mathcal{O}_{3b} . Therefore, the ten points of $\mathcal{O}_{1b} \setminus \mathcal{H}$ partition the set $\mathcal{O}_{3b} \cap \mathcal{H}$, by the distance 2 map, into ten disjoint sets of size sixteen.

Without loss of generality, say $x \in \mathcal{O}_{1b}$. Then the sixteen points of $\mathcal{O}_{3b} \cap \mathcal{H}$ that are at distance 2 from x are at distance 4 from y . Indeed, if $z \in \mathcal{O}_{3b} \cap \mathcal{H}$ lies at distance 2 from x , then through ω , there are precisely five lines containing a point at distance 2 from z . Four of these lines are contained in \mathcal{H} and the fifth line is ωx . So, y which is still on another line through ω should lie at distance 4 from z .

□

We finish this section by proving that the near octagon \mathcal{S} constructed using the valuation geometry of HJ, by taking the points of type A , B , C , and lines of type AAA , ABB , ACC , BBC , CCC , is isomorphic to the near octagon \mathcal{O}_1 constructed using the central involutions of $G_2(4):2$.

Proposition 5.4.15. *The near octagon \mathcal{O}_1 is isomorphic to \mathcal{S} .*

Proof. Regard HJ as a full subgeometry of \mathcal{O}_1 . Then HJ is isometrically embedded into \mathcal{O}_1 by Lemma 5.4.9. By Lemma 2.2.4, every point x of \mathcal{O}_1 will induce a valuation f_x of HJ. This valuation is of Type A if and only if x belongs to HJ. By Lemma 5.4.13, each induced valuation has a unique point with value 0. So, all induced valuations have Type A , B or C by Table 2.7. By Lemma 5.4.14, all induced valuations are distinct, implying that the 4095 induced valuations are precisely the 4095 valuations of HJ that have Type A , B or C . Now, every point of \mathcal{O}_1 is incident with precisely 11 lines. By looking at the columns “A” and “C” of Table 2.8, we see that all lines of \mathcal{V} of Type AAA , ABB , ACC , BBC and CCC should be induced by some line of \mathcal{O}_1 . The number of such lines of \mathcal{S} is equal to $\frac{315 \cdot 5}{3} + 315 \cdot 1 + 315 \cdot 5 + 3150 \cdot 1 + \frac{3150 \cdot 9}{3} = 15015$. Since \mathcal{O}_1 has $\frac{4095 \cdot 11}{3} = 15015$ lines, we see that the lines of \mathcal{S} that are induced are precisely the lines of Type AAA , ABB , ACC , BBC and CCC . We can now conclude that \mathcal{O}_1 and \mathcal{S} are isomorphic. □

5.5. The Automorphism Groups

5.5.1. Automorphism Group of $G_2(4)$ Near Octagon

Lemma 5.5.1. *The group $G_2(4):2$ acts faithfully on points of \mathcal{O}_1 as a group of automorphism of \mathcal{O}_1 , where the action on the point set (i.e. the central involutions) is given by conjugation.*

Proof. Each $g \in G = G_2(4):2$ determines an automorphism of \mathcal{O}_1 : if x is a central involution and $g \in G$, then $x^g = g^{-1}xg$ is again a central involution. We will show that this action is faithful. We have seen that the central involutions generate the group $G' = G_2(4)$. So the centralizer of the set of central involutions is equal to the centralizer of G' in G , denoted by $C_G(G')$. Since $C_G(G')$ is a normal subgroup of G , and since G' is simple, we must have $G' \cap C_G(G') = \{e\}$, where e is the identity of G . So we have $|G'C_G(G')| = |G'| |C_G(G')| \leq |G| = 2|G'|$. Which shows that either $C_G(G')$ is trivial or isomorphic to the cyclic group C_2 . When it is trivial, the action is faithful. So, say it is not. Then we must have that G is equal to the internal direct product of $G' \cong G_2(4)$ and $C_G(G') \cong C_2$, which is a contradiction to the fact that G is a proper semi-direct product of $G_2(4)$ and C_2 . □

Lemma 5.5.2. *Every automorphism θ of \mathcal{O}_1 permutes the elements of the line spread S_1 and thus determines an automorphism of $(S_1, \mathcal{Q}_1) \cong H(4)^D$.*

Proof. The automorphism θ permutes the quads of \mathcal{O}_1 and hence the lines of \mathcal{O}_1 that can be obtained as intersections of two quads (these are precisely the elements of S_1). □

Lemma 5.5.3. *Suppose θ is an automorphism of \mathcal{O}_1 fixing each line of S_1 . Then θ is the identity.*

Proof. Let x be an arbitrary point of \mathcal{O}_1 , $L = \{x, y, z\}$ a line through x not belonging to S_1 and Q the unique quad through L (see Lemma 5.3.9). The lines of S_1 contained in Q determine a spread of Q . Since $Q \cong W(2)$, it is easily seen that L is the unique line of Q meeting the lines L_x , L_y and L_z of Q (see Figure 1.1). From $L_x^\theta = L_x$, $L_y^\theta = L_y$ and $L_z^\theta = L_z$, it then follows that $x^\theta = x$. \square

Proposition 5.5.4. *The full automorphism group of \mathcal{O}_1 is isomorphic to $G_2(4):2$.*

Proof. From Lemma 5.5.1 we see that $G_2(4):2 \leq \text{Aut}(\mathcal{O}_1)$. By Lemmas 5.5.2 and 5.5.3, we have $|\text{Aut}(\mathcal{O}_1)| \leq |\text{Aut}(H(4)^D)| = |G_2(4):2|$, which shows that there is equality. \square

It is possible to give another proof of Proposition 5.5.4 based on the following lemma.

Lemma 5.5.5. *Let H be a subgroup of $G_2(4):2$ isomorphic to $J_2:2$. Then every automorphism θ of $\mathcal{S}_H \cong \text{HJ}$ extends to precisely one automorphism of \mathcal{G} .*

Proof. The action of θ on the point set of \mathcal{S}_H is given by conjugation by a suitable element of $H \cong J_2:2$. This conjugation also determines an automorphism of \mathcal{O}_1 . To show that θ extends to at most one automorphism of \mathcal{G} , we must show that every automorphism φ of \mathcal{O}_1 that fixes each point of \mathcal{S}_H must be trivial. But this is directly implied by Lemma 5.4.14. \square

Since there are 416 Hall-Janko suboctagons (see Lemma 5.4.12), Lemma 5.5.5 implies that $|\text{Aut}(\mathcal{O}_1)| \leq 416 \cdot |\text{Aut}(\text{HJ})| = 416 \cdot |J_2:2| = |G_2(4):2|$. Lemma 5.5.1 then again implies that $\text{Aut}(\mathcal{O}_1) \cong G_2(4):2$. In fact, this reasoning also gives that the automorphism group is transitive on the Hall-Janko suboctagons, but we already knew this in advance as all maximal subgroups isomorphic to $J_2:2$ are conjugate.

5.5.2. Automorphism Group of $L_3(4)$ Near Octagon

In this section we prove that the automorphism group of the $L_3(4)$ near octagon \mathcal{O}_2 is isomorphic to the group $L_3(4):2^2$. Consider the chain $L_3(4) < L_3(4):2^2 < L_3(4):D_{12}$ of groups and their actions on the sets of points and lines of the projective plane $\text{PG}(2, 4)$ (as described in Section 5.2.2). Recall that $L_3(4):D_{12}$ consists of all collineations and correlations of $\text{PG}(2, 4)$. Let Σ denote the set of all central involutions of the groups $L_3(4):2^2$, or equivalently, of $L_3(4)$. Then every element of Σ is a nontrivial elation of $\text{PG}(2, 4)$. From Lemma 5.2.2 it follows that the action of the group $L_3(4):D_{12}$, and hence of the group $L_3(4):2^2$, on the elations via conjugation is faithful. Just as in Lemma 5.5.1, we see that this action of $L_3(4):2^2$ gives us $\text{Inn}(L_3(4):2^2) \cong L_3(4):2^2 \leq \text{Aut}(\mathcal{O}_2)$. We will now show that the size of the automorphism group is at most $4|L_3(4)|$, thus proving equality. Of course, this can be checked in a computer using SageMath, which can easily compute the automorphism group of the collinearity graph of this near octagon (see Lemma 1.3.2), but we prefer to give a computer free proof.

Lemma 5.5.6. *Every automorphism θ of \mathcal{O}_2 permutes the elements of the line spread S_2 and thus determines an automorphism of $(S_2, \mathcal{Q}_2) \cong \text{H}(4, 1)$.*

Proof. The proof is similar to that of Lemma 5.5.2. \square

Lemma 5.5.7. *Let Q be a quad of \mathcal{O}_2 and L_1, \dots, L_5 the five lines of the line spread S_2 which are contained in Q . If θ is an automorphism of \mathcal{O}_2 which fixes each of these lines, then θ fixes each point of Q .*

Proof. Let x be an arbitrary point of Q , and $L = \{x, y, z\}$ a line of Q through x which is not contained in S_2 . Then since Q is isomorphic to $W(2)$, L is the unique line of Q which intersects the lines $L_x, L_y, L_z \in S_2$ of Q . As $L_x^\theta = L_x$, $L_y^\theta = L_y$ and $L_z^\theta = L_z$, we must have $L^\theta = L$ and in particular $x^\theta = x$. \square

Corollary 5.5.8. *If θ is an automorphism of O_2 which fixes every line in S_2 , then θ fixes every point of O_2 . Consequently, the action of $\text{Aut}(O_2)$ on $(S_2, Q_2) \cong H(4, 1)$ is faithful.*

As a result of Corollary 5.5.8, we can look at $\text{Aut}(O_2)$ as a subgroup of $\text{Aut}(H(4, 1)) \cong L_3(4):D_{12}$. Denote the geometry $(S_2, Q_2) \cong H(4, 1)$ by \mathcal{S} . Note that if (x, L) is a flag of $PG(2, 4)$, and hence a point of \mathcal{S} , then the non-trivial elations which have center x and axis L form an element of the line spread S_2 in the near octagon O_2 constructed using these elations. For a point x of $PG(2, 4)$ let C_x denote the set of all (non-trivial) elations with center x , and for a line L let C_L denote the set of all (non-trivial) elations with axis L . Then these sets correspond bijectively to the quads of O_2 , i.e., the lines of \mathcal{S} . What we have shown is that every automorphism of \mathcal{S} which corresponds to either a collineation or a correlation of $PG(2, 4)$, gives rise to at most one automorphism of O_2 , and we have a subgroup $H \cong \text{Aut}(O_2)$ of this group G of all collineations and correlations of $PG(2, 4)$, and the latter has type $L_3(4):D_{12}$. We will show that there exists a subgroup H' of G such that $|H'| = 3$ and $H \cap H' = \{e\}$, thus proving $|H| \leq |G|/|H'| \leq 4|L_3(4)|$ (recall that for any two subgroups H_1, H_2 of a finite group we have $|H_1 H_2| = |H_1||H_2|/|H_1 \cap H_2|$).

Lemma 5.5.9. *Let Q be a quad of O_2 and let L_1, \dots, L_5 be lines of the spread S_2 which are contained in Q . Let θ is an automorphism of O_2 such that $L_1^\theta = L_1$, $L_2^\theta = L_2$, $L_3^\theta = L_4$, $L_4^\theta = L_5$ and $L_5^\theta = L_3$, then θ permutes the points of L_1 according to a cycle of length 3.*

Proof. Since Q is isomorphic to $W(2)$, we know that if $x \in L_1$ and $y \in L_2$ are collinear, then there is a unique $i \in \{3, 4, 5\}$ for which L_i intersects the lines xy . Now for each $i \in \{3, 4, 5\}$, choose the collinear points $x_i \in L_1$ and $y_i \in L_2$ for which the line $M_i = x_i y_i$ intersects the line L_i . Then we must have $M_3^\theta = M_4$, $M_4^\theta = M_5$ and $M_5^\theta = M_3$. Thus, the points x_3, x_4 and x_5 of L_1 are permuted according to a cycle of length 3. \square

Let σ be a nonidentity *homology* of $PG(2, 4)$, i.e., a collineation which fixes all points on a line L (the axis of σ) and all lines through a point $x \notin L$ (the center of σ). Then σ generates a cyclic group of order 3, which is a subgroup H' of G .

Lemma 5.5.10. *Let σ be a homology of $PG(2, 4)$ that has center x and axis L , and let H' be the subgroup of $G \cong L_3(4):D_{12}$ generated by σ . Then, we have $H \cap H' = \{e\}$.*

Proof. Suppose $\bar{\theta} \in H \cap H' \setminus \{e\}$ is induced by an automorphism θ of the near octagon O_2 . Let Q_1 be the quad of O_2 corresponding to the line L and let Q_2 be the quad of O_2 corresponding to a point y on L . Then $Q_1 \cap Q_2$ is the line of S_2 corresponding to the flag (y, L) of $PG(2, 4)$. Since $\bar{\theta}$ is a homology with axis L it fixes every point on L , and therefore θ fixes every line of S_2 contained in the quad Q_1 . So, by Lemma 5.5.7, θ fixes every point in $Q_1 \cap Q_2$. Since $\bar{\theta}$ fixes the point x and y , it fixes the line $M = xy$, and since it's a homology, it permutes the three lines through y other than M and L in a cycle of length 3. This implies that θ fixes two lines of S_2 contained in Q_2 (the line corresponding to (y, M) and the line corresponding to (y, L)), and permutes the remaining three lines in a cycle of length 3. This contradicts Lemma 5.5.9 as θ fixes the line $Q_1 \cap Q_2$ pointwise. \square

We have thus proved the following.

Proposition 5.5.11. *The full automorphism group of O_2 is isomorphic to $L_3(4):2^2$.*

5.6. Distance Regular Graphs of Soicher

In [129], Soicher constructed distance regular graphs Σ , Υ and Δ with intersection arrays $\{416, 315, 64, 1; 1, 32, 315, 416\}$, $\{56, 45, 16, 1; 1, 8, 45, 56\}$ and $\{32, 27, 8, 1; 1, 4, 27, 32\}$, respectively. It is clear from the intersection arrays, that all of these graphs have diameter 4. The graph Σ is a triple cover of the Suzuki graph (in the sense of [129]) and has $\text{Aut}(\Sigma) \cong 3 \cdot \text{Suz}:2$. Under the action of the stabilizer (which is isomorphic to $G_2(4):2$) of its automorphism group with respect to a vertex x , the orbits are equal to $\Sigma_0(x)$, $\Sigma_1(x)$, $\Sigma_2(x)$, $\Sigma_3(x)$ and $\Sigma_4(x)$, with sizes 1, 416, 4095, 832 and 2, respectively (so in particular, the graph Σ is distance transitive). We will show in Section 5.6.1 we can construct the near octagon O_1 using the graph Σ , and that the subgraph of Σ induced on $\Sigma_2(x)$ is isomorphic to the graph defined on the points of O_1 where two points are adjacent when they are at distance 2 and have a unique common neighbor.

The graph Υ is a triple cover of the unique strongly regular graph of parameters $(162, 56, 10, 24)$ and its automorphism group also acts distance transitively on its vertices with the stabilizer with respect to a vertex isomorphic to $L_3(4):2^2$, and suborbits of sizes 1, 56, 315, 112 and 2. In Section 5.6.2, we will construct the near octagon O_1 from this graph. The third graph Δ is simply the second subconstituent of Υ , i.e., the graph induced on $\Upsilon_2(x)$ for any vertex x , and we will show that it is isomorphic to the graph obtained from the points of O_2 by making two points adjacent when they are at distance 2 from each other and have a unique common neighbor. Soicher has recently proved that the graph Δ is the unique distance regular graph with the given intersection array [131].

5.6.1. The First Graph

Let x be a fixed point of Σ . The local graph Σ_x , i.e., the subgraph induced on the vertices set $\Sigma_1(x)$ of vertices adjacent to x , is isomorphic to the well-known $G_2(4)$ -graph, which is a strongly regular graph with parameters $(v, k, \lambda, \mu) = (416, 100, 36, 20)$. Let $G = \text{Aut}(\Sigma)$. Then the stabilizer G_x is isomorphic to the group $G_2(4):2$, and it is the full automorphism group of the local graph of Σ . Recall that O_1 was constructed using the central involutions of the group G . We have computationally verified that every such involution σ (which is a point of O_1) fixes 32 points of the $G_2(4)$ -graph, which we denote by X_σ . Moreover, the elements of X_σ determine the involution σ uniquely, as the group generated by σ is the unique subgroup of $G_2(4):2$ which fixes X_σ pointwise. Therefore, the 4095 points of O_1 are in bijective correspondence with 4095 such *special 32-sets* in the $G_2(4)$ -graph. We also record some information about these special sets in Table 5.3 with respect to the suborbit diagram of O_1 (Figure 5.1).

$\sigma_2 \in \mathcal{O}_0(\sigma_1)$	$X_{\sigma_1} = X_{\sigma_2}$	$\sigma_2 \in \mathcal{O}_{2b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 8$
$\sigma_2 \in \mathcal{O}_{1a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$	$\sigma_2 \in \mathcal{O}_{3a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$
$\sigma_2 \in \mathcal{O}_{1b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 16$	$\sigma_2 \in \mathcal{O}_{3b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 5$
$\sigma_2 \in \mathcal{O}_{2a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$	$\sigma_2 \in \mathcal{O}_4(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 1$

Table 5.3.: Fixed set intersections in the $G_2(4)$ near octagon

In Soicher's graph Σ we can computationally check that for every vertex $y \in \Sigma_2(x)$, the set $\Sigma_1(x) \cap \Sigma_1(y)$ is a *special 32-set* of the $G_2(4)$ -graph induced on $\Sigma_1(x)$, i.e., there exists a unique central involution σ_y of $G_2(4):2$ which fixes this set X_y pointwise. In this manner, we get a map $y \mapsto \sigma_y$ between the 4095 elements of $\Sigma_2(x)$ and the 4095 central involutions of $G_2(4):2$ which is *bijective*. Moreover, computations in the graph Σ give us the following information for two points $y_1, y_2 \in \Sigma_2(x)$ with $S = \Sigma_1(y_1) \cap \Sigma_1(y_2)$, recorded in Table 5.4.

	$d_\Sigma(y_1, y_2)$	$ \Sigma_1(x) \cap S $	$ \Sigma_2(x) \cap S $	$ \Sigma_3(x) \cap S $
$y_2 \in \mathcal{O}_0(y_1)$	0	32	320	64
$y_2 \in \mathcal{O}_{1a}(y_1)$	4	0	0	0
$y_2 \in \mathcal{O}_{1b}(y_1)$	2	16	16	0
$y_2 \in \mathcal{O}_{2a}(y_1)$	2	0	16	16
$y_2 \in \mathcal{O}_{2b}(y_1)$	1	8	76	16
$y_2 \in \mathcal{O}_{3a}(y_1)$	3	0	0	0
$y_2 \in \mathcal{O}_{3b}(y_1)$	2	5	25	2
$y_2 \in \mathcal{O}_4(y_1)$	2	1	25	6

Table 5.4.: Intersection patterns in Soicher's first graph

From Tables 5.3 and 5.4, it follows that for two vertices y_1, y_2 in $\Sigma_2(x)$, the involutions σ_{y_1} and σ_{y_2} are collinear in the near octagon \mathcal{O}_1 (which is equivalent to $\sigma_2 \in \mathcal{O}_{1a}(\sigma_1) \cup \mathcal{O}_{1b}(\sigma_1)$) if and only if $d_\Sigma(y_1, y_2) = 4$ or $|\Sigma_1(x) \cap \Sigma_1(y_1) \cap \Sigma_1(y_2)| = 16$ (and thus $d_\Sigma(y_1, y_2) = 2$). From the discussion so far, we have the following.

Theorem 5.6.1. *Let x be a fixed vertex of Σ .*

- (1) *Let Γ denote the graph defined on the set $\Sigma_2(x)$ of vertices at distance 2 from x in Σ , by making two vertices y_1, y_2 adjacent if and only if $d_\Sigma(y_1, y_2) = 4$ or $|\Sigma_1(x) \cap \Sigma_1(y_1) \cap \Sigma_1(y_2)| = 16$. Then Γ is isomorphic to the collinearity graph of the $G_2(4)$ near octagon \mathcal{O}_1 .*
- (2) *The subgraph of Σ induced on the vertex set $\Sigma_2(x)$ is isomorphic to the graph defined on the points of the $G_2(4)$ near octagon by making two points adjacent when they are at distance 2 and have a unique common neighbor.*

5.6.2. The Second Graph

A similar analysis can be done for the second graph Υ . Its local graph is isomorphic to the Gewirtz graph, which is the unique strongly regular graph with parameters $(56, 10, 0, 2)$ [31]. Again, the stabilizer at a vertex of Υ is isomorphic to $L_3(4):2^2$, which is also the automorphism group of the Gewirtz graph. For every central involution σ of $L_3(4):2^2$, there is *special 8-set* X_σ consisting of the points of Gewirtz graph that are left fixed by this involution. The map $\sigma \mapsto X_\sigma$ is a bijection between the 315 vertices of the $L_3(4)$ near octagon \mathcal{O}_2 and the 315 special 8-sets in the Gewirtz graph. Fix a vertex x in the graph Υ . Then it can be checked computationally that for every vertex $y \in \Upsilon_2(x)$, the set $\Upsilon_1(x) \cap \Upsilon_1(y)$ is a special 8-set of the local graph at x . The information regarding the special sets and intersection patterns in this case is given in Tables 5.5 and 5.6, from which the following result follows.

$\sigma_2 \in \mathcal{O}_0(\sigma_1)$	$X_{\sigma_1} = X_{\sigma_2}$	$\sigma_2 \in \mathcal{O}_{2b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 2$
$\sigma_2 \in \mathcal{O}_{1a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$	$\sigma_2 \in \mathcal{O}_{3a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$
$\sigma_2 \in \mathcal{O}_{1b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 4$	$\sigma_2 \in \mathcal{O}_{3b}(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 2$
$\sigma_2 \in \mathcal{O}_{2a}(\sigma_1)$	$X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$	$\sigma_2 \in \mathcal{O}_4(\sigma_1)$	$ X_{\sigma_1} \cap X_{\sigma_2} = 1$

Table 5.5.: Fixed set intersections in $L_3(4)$ near octagon

	$d_{\Upsilon}(y_1, y_2)$	$ \Upsilon_1(x) \cap S $	$ \Upsilon_2(x) \cap S $	$ \Upsilon_3(x) \cap S $
$y_2 \in \mathcal{O}_0(y_1)$	0	8	32	16
$y_2 \in \mathcal{O}_{1a}(y_1)$	4	0	0	0
$y_2 \in \mathcal{O}_{1b}(y_1)$	2	4	4	0
$y_2 \in \mathcal{O}_{2a}(y_1)$	2	0	4	4
$y_2 \in \mathcal{O}_{2b}(y_1)$	1	2	4	4
$y_2 \in \mathcal{O}_{3a}(y_1)$	3	0	0	0
$y_2 \in \mathcal{O}_{3b}(y_1)$	2	2	4	2
$y_2 \in \mathcal{O}_4(y_1)$	2	1	4	3

Table 5.6.: Intersection patterns in Soicher's second graph

Theorem 5.6.2. *Let x a fixed vertex of Υ .*

- (1) *Let Γ denote the graph defined on the set $\Upsilon_2(x)$ of vertices at distance 2 from x in Υ , by making two vertices y_1, y_2 adjacent if and only if $d_{\Upsilon}(y_1, y_2) = 4$ or $|\Upsilon_1(x) \cap \Upsilon_1(y_1) \cap \Upsilon_1(y_2)| = 4$. Then Γ is isomorphic to the collinearity graph of the $L_3(4)$ near octagon \mathcal{O}_2 .*
- (2) *The subgraph of Υ induced on the vertex set $\Upsilon_2(x)$ (which is the third distance regular graph of Soicher, Δ) is isomorphic to the graph defined on the points of the $L_3(4)$ near octagon by making two points adjacent when they are at distance 2 and have a unique common neighbor.*

6. Suzuki Tower of Near Polygons

6.1. Introduction

The Suzuki tower¹ is the sequence of five finite simple groups $L_3(2) < U_3(3) < J_2 < G_2(4) < Suz$ where each group, except the last one, is a maximal subgroup of the next group in the sequence. The five groups in the Suzuki tower correspond to five vertex-transitive graphs $\Gamma_0, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ where the automorphism groups of the Γ_i 's are $L_3(2):2, U_3(3):2, J_2:2, G_2(4):2$ and $Suz:2$, respectively [132], [140, Section 5.6]. The graph Γ_0 is the complement of the incidence graph of Fano plane, i.e., the quartic vertex transitive co-Heawood graph, and the rest of them are strongly regular graphs having the following parameters (v, k, λ, μ) :

- Γ_1 : $(36, 14, 4, 6)$, the $U_3(3)$ -graph;
- Γ_2 : $(100, 36, 14, 12)$, the Hall-Janko graph;
- Γ_3 : $(416, 100, 36, 20)$, the $G_2(4)$ -graph;
- Γ_4 : $(1782, 416, 100, 96)$, the Suzuki graph.

With the near octagon O_1 that we constructed in Chapter 5, we now also have a “tower” of near polygons

$$H(2, 1) \subset H(2)^D \subset HJ \subset O_1,$$

with automorphism groups $L_3(2):2, U_3(3):2, J_2:2$ and $G_2(4):2$, respectively. Thus we can call this the “Suzuki tower of near polygons”. Moreover, we will see in this chapter that the five graphs $\Gamma_0, \dots, \Gamma_4$ can be constructed in a uniform fashion using these four near octagons, as follows: Let $\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$ be the Suzuki tower near polygons and $\Gamma_0, \Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ the graphs of the Suzuki tower. We define \mathcal{S}_{-1} to be the partial linear space on nine points and four lines obtained from the (3×3) -grid by removing two disjoint lines (and keeping the points incident with these two lines)

Theorem 6.1.1. *The complement of Γ_i with $i \in \{0, 1, 2, 3\}$ is isomorphic to the graph whose vertices are the subgeometries of \mathcal{S}_i isomorphic to \mathcal{S}_{i-1} , where two distinct subgeometries are adjacent whenever they intersect in the perp of a point.*

For $i \in \{0, 1, 2\}$ the statement of Theorem 6.1.1 was already known in the literature (thought without any mention of the Suzuki tower of near polygons), but for $i = 3$ our result is new. While we do not give a construction of Γ_4 in a similar fashion, we can translate the original construction by Suzuki [132] in near polygon language as follows.

Theorem 6.1.2. *Define a graph as follows:*

¹a name coined by Tits, as it has been mentioned in [101]

- take the elements of $\{\infty\}$, A and B as vertices, where ∞ is just a symbol, A is the set of all Hall-Janko suboctagons of the near octagon O_1 and B is the line spread S_1 of O_1 ;
- join ∞ to all vertices in A , join two distinct vertices of A if the corresponding suboctagons intersect in a subhexagon isomorphic to $H(2)^D$, join a vertex of A to all the vertices in B that correspond to a line intersecting the suboctagon, join two vertices in B if the corresponding lines are at distance 2 from each other in the near octagon.

Then this graph is isomorphic to the Suzuki graph.

We will prove Theorems 6.1.1 and 6.1.2 in Section 6.2.

Remark. (1) We note that there are various other geometries associated with the Suzuki tower that have been studied before, see for example [112], [128] and [101].

- (2) We do not know if there is a near polygon which corresponds to Suz , but certainly the involution geometry of Suz studied in the literature [16, 141] is not a near polygon; we can directly see from the suborbit diagram [16, Fig. 1] that there are point-line pairs (p, L) in this geometry where every point of the line L is at the same distance 4 from p . However, this involution geometry is a near 9-gon in the sense of [34, Sec. 6.4]. Indeed, it is clear from the suborbit diagram [16, Fig. 1] that the geometry has diameter 4 and that for every point-line pair (p, L) with $d(p, L) < 4$, there is a unique point on L nearest to p . Similarly, one can verify that the involution geometry of the Conway group Co_1 [17, Fig. 1] which contains Suz is a near 11-gon. The techniques involving valuations of near polygons used in this thesis do not work for near $(2d+1)$ -gons, and thus we are unable to give similar characterizations of these involution geometries. We do not know if there are more near polygons hiding in larger groups that can extend the Suzuki tower of near polygons (see [128] for a possible extension of the Suzuki tower of groups).

For all $i \in \{1, 2, 3, 4\}$ the graph Γ_i contains Γ_{i-1} as a local graph, i.e., the graph induced on the neighborhood of any vertex of Γ_i is isomorphic to Γ_{i-1} . In fact, it has been proved in [115] that for $i \in \{2, 3\}$ the graph Γ_i is the unique connected graph which is locally Γ_{i-1} , while there exist two connected graphs which are locally Γ_3 , the graph Γ_4 and the graph $3\Gamma_4$ constructed by Soicher [129] which is a 3-fold antipodal cover of the Suzuki graph Γ_4 . In [36], it was proved that there are up to isomorphism three connected graphs that are locally Γ_0 , one of which is the $U_3(3)$ -graph Γ_1 . Motivated by these characterizations of the graphs in the Suzuki tower, we give some characterizations of the geometries in the Suzuki tower of near polygons by proving the following three theorems.

Theorem 6.1.3. *The dual split Cayley hexagon of order 2 is the unique near hexagon of order $(2, 2)$ that contains the generalized hexagon $H(2, 1)$ as an isometrically embedded subgeometry.*

Theorem 6.1.4. *The Hall-Janko near octagon is the unique near octagon of order $(2, 4)$ that contains the dual split Cayley hexagon of order $(2, 2)$ as an isometrically embedded subgeometry.*

Theorem 6.1.5. *The $G_2(4)$ near octagon is the unique near octagon of order $(2, 10)$ that contains the Hall-Janko near octagon as an isometrically embedded subgeometry.*

The proof of Theorem 6.1.3 already follows from the results of Section 1.3.2. It is folklore that $H(2)$ does not contain any subhexagons isomorphic to $H(2, 1)$.² Therefore, it suffices to show that $\mathbb{L}_3 \times \mathbb{L}_3 \times \mathbb{L}_3$ does not contain any subgeometry isomorphic to $H(2, 1)$. So, let $\mathcal{H}_1 \cong H(2, 1)$ be a subgeometry of $\mathcal{H}_2 \cong \mathbb{L}_3 \times \mathbb{L}_3 \times \mathbb{L}_3$. We know that \mathcal{H}_1 has 21 points and \mathcal{H}_2 has 27 points. Each point of \mathcal{H}_1 is collinear with exactly two points of $\mathcal{H}_2 \setminus \mathcal{H}_1$. Since $|\mathcal{H}_2 \setminus \mathcal{H}_1| = 6$, there must be a point in $\mathcal{H}_2 \setminus \mathcal{H}_1$ collinear with at least $(21 \cdot 2)/6 = 7$ points of \mathcal{H}_1 . This contradicts the fact that \mathcal{H}_2 has order $(2, 2)$, and thus we have proved Theorem 6.1.3. We will prove Theorems 6.1.4 and 6.1.5 in Section 6.3 using the theory of valuations described in Chapter 2.

6.2. Constructing the Graphs

We first review the original construction of the Suzuki Tower graphs [132] (see also [115]). The graph Γ_0 is simply the complement of the incidence graph of the Fano plane $PG(2, 2)$, which clearly has automorphism group $L_3(2):2$. Let $\Delta = \Gamma_{i-1}$, $\Gamma = \Gamma_i$ and $H = \text{Aut}(\Delta)$ for some $i \in \{1, 2, 3, 4\}$. The graph Γ is then constructed from the graph Δ as follows.

Let S be the conjugacy class of 2-subgroups (involutions) of H of type $2A$ (in ATLAS notation) if $i \in \{1, 2, 3\}$. In the remaining case, $i = 4$, let S be the conjugacy class of 2^2 -subgroups of H generated by $2A$ -involutions x and y which satisfy the condition $[H : N_H(\langle x, y \rangle)] = 1365$. Let ∞ be an extra symbol, which is not contained in S or the set of vertices $V(\Delta)$ of the graph Δ . Then the vertex set of Γ is $V(\Gamma) = \{\infty\} \cup V(\Delta) \cup S$, and adjacencies are defined as:

- (a) the vertex ∞ is made adjacent to all the elements of $V(\Delta)$;
- (b) two vertices in $V(\Delta)$ are adjacent if they are adjacent as vertices of Δ ;
- (c) a vertex $x \in S$ is adjacent to a vertex $v \in V(\Delta)$ if a non trivial element of the subgroup corresponding to x fixes v ;
- (d) two vertices x, y in S are adjacent if x, y considered as subgroups of H do not commute but there exists a $z \in S$ that commutes with both of them.

While the construction seems “asymmetric”, the group $\text{Aut}(\Gamma)$ acts transitively on the vertices. Also, for each $i \in \{1, 2, 3, 4\}$, the group $G_{i-1} = \text{Aut}(\Gamma_{i-1})$ is the stabilizer of the group $G_i = \text{Aut}(\Gamma_i)$ at a vertex. Moreover, G_{i-1} is a maximal subgroup of G_i of index $|\Gamma_i|$. We have $G_0 \cong L_3(2):2$, $G_1 \cong G_2(2)$, $G_2 \cong J_2:2$, $G_3 \cong G_2(4):2$ and $G_4 \cong Suz:2$.

In the construction of the near octagon O_1 in Chapter 5 we saw that the 2^2 -subgroups of $G = G_2(4):2$ which are generated by $2A$ involutions x, y satisfying $[G : N_G(\langle x, y \rangle)] = 1365$, correspond (bijectively) to the lines of O_1 which are contained in the line spread S_1 . Moreover, in the $G_2(4)$ graph we can check that two vertices are adjacent if and only if the corresponding $J_2:2$ subgroups (stabilizers of the vertices) intersect in a $G_2(2)$ subgroup. Also recall the correspondence between J_2 subgroups of $G_2(4)$ and the $J_2:2$ subgroups of $G_2(4):2$ described in Section 5.4. Therefore, the $G_2(4)$ -graph is the graph whose vertices are the maximal subgroup of $G_2(4)$ isomorphic to J_2 where two Hall-Janko subgroups are adjacent whenever they intersect in a subgroup isomorphic to $G_2(2)' \cong U_3(3)$.

²A proof of this can be given by checking that if $H(1, 2)$ is contained in $H(2)^D$ then one can get an intersecting set in the sense of [120] of size 1, which contradicts [120, Corollary 1.2].

Lemma 6.2.1. *The $G_2(4)$ -graph Γ_3 is isomorphic to the graph Γ' whose vertices are the Hall-Janko suboctagons of O_1 , where two Hall-Janko suboctagons are adjacent whenever they intersect in a subgeometry isomorphic to $H(2)^D$.*

Proof. It is well-known that the Hall-Janko near octagon HJ has 100 subhexagons isomorphic to $H(2)^D$ (see for example [30]), and that these are in bijective correspondence with the 100 maximal subgroups of J_2 isomorphic to $G_2(2)'$ (just like the J_2 subgroups of $G_2(4)$ correspond to the Hall-Janko near octagons of the $G_2(4)$ near octagon). The points of a subhexagon are the central involutions contained in the corresponding maximal subgroup. Moreover, these central involutions generate the maximal subgroup.

For every subgroup H of $G_2(4)$, denote by Σ_H the set of central involutions which are contained in H . If $H \cong J_2$, then the geometry \mathcal{S}_H induced on the subspace Σ_H of O_1 is isomorphic to HJ. By Lemma 5.4.12, the map $H \mapsto \mathcal{S}_H$ defines a bijection between the 416 maximal subgroups of $G_2(4)$ isomorphic to J_2 and the 416 Hall-Janko suboctagons of O_1 . We show that this map defines an isomorphism between the $G_2(4)$ -graph and the graph Γ' .

Take two mutually distinct subgroups H_1 and H_2 of $G_2(4)$ isomorphic to J_2 . If H_1 and H_2 are two adjacent vertices of the $G_2(4)$ -graph, then the subgeometries \mathcal{S}_{H_1} and \mathcal{S}_{H_2} intersect in a subgeometry whose point set is $\Sigma_{H_1} \cap \Sigma_{H_2} = \Sigma_{H_1 \cap H_2}$, i.e. in a subgeometry isomorphic to $H(2)^D$ as $H_1 \cap H_2 \cong G_2(2)'$. Conversely, suppose that \mathcal{S}_{H_1} and \mathcal{S}_{H_2} intersect in a subgeometry isomorphic to $H(2)^D$. Then $\Sigma_{H_1} \cap \Sigma_{H_2}$ contains all central involutions that are contained in a certain $G_2(2)'$ -subgroup K_i of $H_i \cong J_2$, $i = 1, 2$. Since all these central involutions generate K_i , the groups K_1 and K_2 are equal, say to K . As K is a maximal subgroup of both H_1 and H_2 , we have $K = H_1 \cap H_2$, i.e. H_1 and H_2 are adjacent in the $G_2(4)$ -graph. \square

Therefore, we see that the $G_2(4)$ graph has a symmetric construction using the Hall-Janko suboctagons of the new near octagon O_1 . It can be easily checked in the computer model that when two such suboctagons do not intersect in a subgeometry isomorphic to $H(2)^D$, they intersect in a point pencil (set of point at distance at most 1 from a fixed point) of suboctagons. And thus we obtain Theorem 6.1.1 for the case of $i = 3$. The remaining cases of $i \in \{0, 1, 2\}$ also follow from similar reasoning, but those results are “well-known” and we skip the details here. Turning our focus to the Suzuki graph, we see that as a consequence of Lemma 6.2.1 the construction given in Theorem 6.1.2 is the geometrical translation of the original construction given by Suzuki. Note that two lines of the lines spread S_1 are at distance 1 from each other if and only if the corresponding 2^2 -subgroups of $G_2(4):2$ commute. The only thing left to check is the adjacency between elements of $V(\Delta)$ and S , where Δ is the $G_2(4)$ -graph and S is the line spread S_1 of O_1 . From Lemma 5.5.5 we know that every automorphism of O_1 which stabilizes a Hall-Janko suboctagon \mathcal{H} must be a conjugation by an element of the $J_2:2$ -subgroup corresponding to \mathcal{H} . From this it follows that if $x \in S$ and $v \in V(\Delta)$, then a non-trivial element of the subgroup corresponding to x fixes v if and only if the spread line corresponding to x intersects the Hall-Janko suboctagon corresponding to v .

6.3. Characterization of the Suzuki Tower

6.3.1. Hall-Janko Near Octagon

For this section let \mathcal{S}' be a near octagon of order $(2, 4)$ with a generalized hexagon \mathcal{S} isomorphic to $H(2)^D$ isometrically embedded in it. The valuation geometry \mathcal{V} of \mathcal{S} is described in Tables 2.3 and 2.4. From Table 2.4, it can be seen that the set of valuations of type A and B form a subspace of \mathcal{V} , and hence we can define a full subgeometry $\mathcal{V}_{A,B}$ of \mathcal{V} induced by these valuations. In this section we will show that \mathcal{S}' is isomorphic to $\mathcal{V}_{A,B}$. Since HJ is a near octagon of order $(2, 4)$ with $H(2)^D$ isometrically embedded in it, this will prove Theorem 6.1.4.

Lemma 6.3.1. (1) *Each point of \mathcal{S}' is at distance at most 2 from \mathcal{S} .*

(2) *Points at distance 1 from \mathcal{S} must be of type A , B or C .*

(3) *Points at distance 2 from \mathcal{S} must be of type C or D .*

Proof. Since the maximum value of a valuation of \mathcal{S} is at least 2 (see Table 2.3) and the diameter of \mathcal{S}' is 4, by Lemma 2.2.4, the distance of any point of \mathcal{S}' to \mathcal{S} is at most 2. This proves (1). Again by Lemma 2.2.4, if a point x of \mathcal{S}' lies at distance 2 from \mathcal{S} , then f_x has maximum value at most 2, implying that x can only be of type C or D .

Now, let x be a point of type D at distance 1 from \mathcal{S} . The five points with f_x -value 0 in \mathcal{S} must be collinear with x and necessarily be of type A (all points in \mathcal{S} are of type A). By Theorem 2.3.2 and Table 2.4 this gives rise to five distinct \mathcal{V} -lines of type ADD through a valuation of type D in the valuation geometry \mathcal{V} of \mathcal{S} , which contradicts the corresponding entry in Table 2.4. \square

Lemma 6.3.2. *Each point of \mathcal{S}' at distance 1 from \mathcal{S} must be of type A or B .*

Proof. Let x be a point of type C at distance 1 from \mathcal{S} . We see from Table 2.3 that there is a unique point x' in \mathcal{S} collinear with x . Again from Table 2.3 we see that there are 22 points with f_x -value 1 in \mathcal{S} , which must necessarily be at distance 2 from x . Six of these points are neighbors of x' in \mathcal{S} and these are the only ones that have a common neighbor with x that lies inside \mathcal{S} (namely x'). The remaining 16 points give rise to neighbors of x that lie outside \mathcal{S} . Since the order of \mathcal{S}' is $(2, 4)$ and x' lies in \mathcal{S} there are only 9 neighbors of x that lie outside \mathcal{S} . Therefore, at least one such neighbor y must be collinear with more than one point in \mathcal{S} . By Lemma 6.3.1 the point y must be of type A , B or C and for each of these possibilities we have $|\mathcal{O}_{f_y}| = 1$. This is a contradiction. \square

Lemma 6.3.3. *If x, y are two points of \mathcal{S}' , not contained in \mathcal{S} , of type A and B respectively, then x and y cannot be collinear.*

Proof. Let x, y be such points and suppose they are collinear. By Lemma 6.3.1 they must be at distance 1 from \mathcal{S} . The three valuations induced by the three points on the line xy must be distinct (see Theorem 2.3.2) and therefore, the line xy gives rise to a \mathcal{V} -line of $\mathcal{S} \cong H(2)^D$. From Table 2.4 it follows that the line xy is of type ABB . Let y' be the unique neighbor of y in \mathcal{S} . Then by a similar reasoning the line yy' is also of type ABB . But, in the valuation geometry there is a unique line of type ABB through a valuation of type B . Therefore, the points x and y' induce the same type A valuation, which shows that $\mathcal{O}_{f_x} = \{y'\}$ and hence, x and y' are collinear. This contradicts the fact that \mathcal{S}' is a near polygon. \square

Lemma 6.3.4. *If x is a point of type B in \mathcal{S}' , then it has a unique neighbor in \mathcal{S} and all the other neighbors of x must induce distinct type B valuations of \mathcal{S} .*

Proof. Let x be such a point, necessarily at distance 1 from \mathcal{S} by Lemma 6.3.1. By Lemma 2.2.4 and Table 2.3, it has a unique neighbor, say x' , in \mathcal{S} . There are 14 points with f_x -value 1 in \mathcal{S} and 6 of them are the neighbors of x' . The remaining 8 must give rise to neighbors of x lying outside \mathcal{S} . Let y be such a neighbor. By Lemmas 6.3.2 and 6.3.3, y must be of type B and then by Lemma 2.2.4 it cannot lie on the line xx' (as then y , which is a type B valuation, will have 2 neighbors in \mathcal{S}). Therefore, we get 8 type B neighbors of x each corresponding to a distinct valuation of \mathcal{S} (since the set \mathcal{O}_f is distinct for each such valuation). The third point on the line xx' must also be of type B and induce a valuation distinct from all other type B neighbors of x . Since the order of \mathcal{S}' is $(2, 4)$, we have accounted for all neighbors of x . \square

The following is an immediate consequence of Lemma 6.3.4.

Corollary 6.3.5. *There are no lines in \mathcal{S}' of type BCC or BDD .*

Lemma 6.3.6. *There is no point in \mathcal{S}' of type C or D .*

Proof. Let x be such a point, necessarily at distance 2 from \mathcal{S} (see Lemmas 6.3.1 and 6.3.2). We treat the two cases separately.

Case 1: Let x be of type D . By Table 2.3, $|\mathcal{S} \cap \Gamma_2(x)| = 5$. Every line through x that contains a point in $\Gamma_1(\mathcal{S})$ must be of type ADD by Table 2.6, Lemma 6.3.2 and Corollary 6.3.5. Since each of these lines has exactly one point which lies in $\Gamma_1(\mathcal{S})$, and since that point (of type A) has a unique neighbor in \mathcal{S} , it must be the case that all five lines through x are of type ADD . In fact, this also shows that these five lines correspond to five distinct lines in the valuation geometry. But we know from Table 2.4 that there are only three lines of type ADD through a point of type D in the valuation geometry, a contradiction.

Case 2: Let x be of type C . Since $|\mathcal{O}_{f_x}| = 1$, we have $\mathcal{O}_{f_x} = \Gamma_2(x) \cap \mathcal{S} = \{x'\}$ for some $x' \in \mathcal{S}$. Since there are no points of type D (by *Case 1*) and no lines of type BCC (by Corollary 6.3.5), all lines through x must be of type ACC or CCC by Table 2.4. Each of the type A neighbors of x induces the valuation $f_{x'}$ of \mathcal{S} . Therefore, besides the 6 neighbors of x' in \mathcal{S} , every point of \mathcal{S} that has f_x -value 1 must be at distance 2 from a type C neighbor of x . There are $16 = 22 - 6$ points with f_x -value 1 in \mathcal{S} that are not neighbor of x' . Since there are at most 9 type C neighbors of x , there must be a type C neighbor y of x which is at distance 2 from two distinct points of \mathcal{S} . This contradicts the fact that $|\mathcal{O}_{f_y}| = 1$ (y is a type C point). \square

By Lemmas 6.3.1 and 6.3.6, we have:

Corollary 6.3.7. *Every point x of \mathcal{S}' not contained in \mathcal{S} has type A or B , and lies at distance 1 from a unique point of \mathcal{S} (which we will refer to as the projection of x in \mathcal{S}).*

Lemma 6.3.8. *Let Q be a quad of \mathcal{S}' that intersects \mathcal{S} nontrivially. Then $Q \cap \mathcal{S}$ is either a singleton or a line.*

Proof. Say $Q \cap \mathcal{S}$ is not a singleton. Since $Q \cap \mathcal{S}$ is a subspace of \mathcal{S}' , it suffices to show that there are no two non-collinear points in $Q \cap \mathcal{S}$. Let x, y be two non-collinear points in $Q \cap \mathcal{S}$. Since Q is a non-degenerate generalized quadrangle, there are at least two common neighbors of x and y in Q . Since points at distance 2 in $\mathcal{S} \cong H(2)^D$ have a unique common neighbor and $Q \cap \mathcal{S}$ is a convex subspace, at least one of these common neighbors must lie outside \mathcal{S} . This gives rise to a point at distance 1 from \mathcal{S} with two neighbors (x and y) in \mathcal{S} , which contradicts Corollary 6.3.7. \square

Lemma 6.3.9. *If x is a point of type A in \mathcal{S}' which is not contained in \mathcal{S} , then there exists a unique $W(2)$ -quad Q containing x and its projection x' in \mathcal{S} . For this quad Q we have:*

- (a) Q intersects \mathcal{S} in a line M .
- (b) Let M_1 and M_2 be the two lines through x' in \mathcal{S} other than M . Then the two lines through x that are not contained in Q can be labelled L_1 and L_2 such that L_1, M_1 are parallel and at distance 1 from each other, and L_2, M_2 are parallel and at distance 1 from each other.

Proof. Let x be a point of type A outside \mathcal{S} and let x' be the unique point in \mathcal{S} collinear with x . We have $f_x = f_{x'}$. Each of the four lines through x which lie outside \mathcal{S} are of type AAA by Lemmas 6.3.3 and 6.3.6. From Table 2.4 we see that there are only three distinct lines of type AAA through a valuation of type A in the valuation geometry. Therefore, there exists two lines K_1, K_2 through x which lie in $\Gamma_1(\mathcal{S})$ and induce the same set of valuations on \mathcal{S} . Let $\{f_1, f_2, f_3\}$ be this set with $f_1 = f_x = f_{x'}$. Since all three lines through x' which lie inside \mathcal{S} correspond to distinct lines of type AAA in the valuation geometry, at least one of them, say M , must induce the set $\{f_1, f_2, f_3\}$ of valuations.

Let $y \neq x'$ be a point on M . Then y is collinear with a point on K_1 and a point on K_2 both of which induce the valuation equal to f_y . Therefore, x, y are two points at distance 2 in a near polygon with at least three common neighbors. From Theorem 1.3.8, it follows that x and y lie in a unique quad Q . By the classification of quads of order $(2, t)$ and Lemma 1.3.13, Q must be isomorphic to $W(2)$, the unique generalized quadrangle of order $(2, 2)$. From Lemma 6.3.8 it follows that $Q \cap \mathcal{S} = M$ and hence, for each point on M , the two lines through it going out of \mathcal{S} are contained in Q . Hence none of the points on M can be contained in a $W(2)$ -quad other than Q .

Now let L be a line through x not contained in Q . L necessarily induces a set of valuations other than $\{f_1, f_2, f_3\}$. There are only two other possibilities and both of them are induced by lines through x' contained in \mathcal{S} , but not in Q . Therefore there must be a line L' through x' inducing the same set of valuations as L . The correspondence $L \mapsto L'$ between the set of lines through x not contained in Q and the set of lines through x' in \mathcal{S} distinct from M is a bijection as otherwise there would exist another $W(2)$ -quad through the line xx' but we have already proved that there is a unique such quad. \square

Lemma 6.3.10. *There is no point in \mathcal{S}' of type A outside \mathcal{S} .*

Proof. Let x be a point of type A outside \mathcal{S} . By Lemma 6.3.9 it lies in a unique $W(2)$ -quad Q which intersects \mathcal{S} in a line L . By Lemma 6.3.3 and Corollary 6.3.7, all points of $Q \setminus L$ have type A . Let x' be the projection of x in \mathcal{S} and y' a neighbor of x' in \mathcal{S} lying on a line through x' other than L . By Lemma 6.3.9, there exists a unique neighbor y of y' outside \mathcal{S} and collinear with x . Again by Lemma 6.3.3 and Corollary 6.3.7, the point y has type A . So, by Lemma 6.3.9, there exists a unique $W(2)$ -quad S containing y and y' . The $W(2)$ -quads Q and S are disjoint.

Suppose p is a neighbor of x' contained in $Q \setminus L$. As p has type A , there exists by Lemma 6.3.9 a unique line through p disjoint from \mathcal{S} that is parallel and at distance 1 from the line $x'y'$, implying that there is a common neighbor of p and y' in $\mathcal{S} \setminus \mathcal{S}$. This implies that we can label the two lines of Q through x' distinct from $L = Q \cap \mathcal{S}$ by T_1 and T_2 and the two lines of S through y' distinct from $S \cap \mathcal{S}$ by U_1 and U_2 such that T_1, U_1

are parallel and at distance 1 from each other, and T_2, U_2 are parallel and at distance 1 from each other.

Now, consider a point z in S which is not collinear with y' . If z is at distance 1 from Q (necessarily from a point of type A of $Q \setminus L$), then by Lemma 6.3.9 its projection $z' \in S \cap \mathcal{S}$ is collinear with a point on the line L in \mathcal{S} , contradicting the fact that \mathcal{S} is a generalized hexagon. So z (as well as every point of S non-collinear with y') must be at distance at least 2 from Q . The two lines of S through y' distinct from $S \cap \mathcal{S}$ are parallel and at distance 1 from a line of Q . So, taking the projection of z on these two lines, we see that there are two points in Q at distance 2 from z . By Theorem 2.2.2, the point z induces a classical or an ovoidal valuation of Q . Since there are two points in Q at distance 2 from z , the point z must induce an ovoidal valuation of Q . Since there are five points in an ovoid in Q ($\cong W(2)$), each of the five lines through z must contain a (necessarily unique) point at distance 1 from Q . Thus the projection of z on \mathcal{S} , z' , must be collinear with a point in Q . Now, all the lines through z' are either in \mathcal{S} or in S . Since S and Q are disjoint and \mathcal{S} is a generalized hexagon, we have a contradiction. \square

Therefore, \mathcal{S}' has the following description:

- each point of \mathcal{S}' is at distance at most 1 from \mathcal{S} ;
- each point of \mathcal{S}' that lies in \mathcal{S} induces a valuation of type A, and
- each point of \mathcal{S}' that does not lie in \mathcal{S} induces a valuation of type B.

Obviously, distinct points of \mathcal{S}' induce distinct type A valuations since their zero sets are distinct. To prove that \mathcal{S}' is isomorphic to $\mathcal{V}_{A,B}$, we first show that no two points in \mathcal{S}' can induce the same type B valuation of \mathcal{S} , which will give us a bijection between the point sets of these geometries.

Lemma 6.3.11. *If g_1 and g_2 are two distinct valuations of type B collinear to each other in the valuation geometry of \mathcal{S} , then we have $|\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_1\}| = |\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_2\}|$*

Proof. Let g_1, g_2 be two such valuations of type B. Say a point y in \mathcal{S}' induces the valuation g_1 of \mathcal{S} . If g_1 and g_2 lie on \mathcal{V} -line of type ABB , then the third point on the line joining y and the unique neighbor of y that lies in \mathcal{S} induces the valuation g_2 , giving us a bijection between $\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_1\}$ and $\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_2\}$. So, say g_1 and g_2 lie on \mathcal{V} -line of type BBB . Then by Lemma 6.3.4 we know that each of the four lines through y which do not intersect \mathcal{S} must induce distinct \mathcal{V} -lines of type BBB . But we know from Table 2.4 that there are exactly four such \mathcal{V} -lines containing g_1 , and hence g_2 is contained in exactly one of them. Therefore, there must be precisely one neighbor of y in $\mathcal{S}' \setminus \mathcal{S}$ which induces the valuation g_2 . This gives a bijection between $\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_1\}$ and $\{x \in \mathcal{S}' \setminus \mathcal{S} \mid f_x = g_2\}$. \square

Lemma 6.3.12. *The subgeometry of \mathcal{V} defined on the type B valuations by the lines of type ABB and BBB is connected.*

Proof. This is easily checked via computations in the computer model of the valuation geometry of $H(2)^D$ obtained using the code given in Chapter 3. \square

Corollary 6.3.13. *For each type B valuation f of \mathcal{S} , there exists exactly one point $x \in \mathcal{S}'$ with $f_x = f$.*

Proof. Since each point of \mathcal{S} is collinear with precisely four points of $\mathcal{S}' \setminus \mathcal{S}$, and each point of $\mathcal{S}' \setminus \mathcal{S}$ has a unique neighbor in \mathcal{S} , we have $|\mathcal{S}' \setminus \mathcal{S}| = 4|\mathcal{S}| = 252$. By Lemmas 6.3.11 and 6.3.12 we know that for every pair of type B valuations there exist equally many points in \mathcal{S}' which induce those valuations. But from Table 2.3 we can see that there are exactly 252 valuations of type B . Therefore each type B valuation is induced exactly once. \square

Now we can prove that \mathcal{S}' is isomorphic to $\mathcal{V}_{A,B}$ as follows. Map every point x of \mathcal{S}' to the valuation of type $T \in \{A, B\}$ that it induces. Since no two points of \mathcal{S}' induce the same valuation (see Corollary 6.3.13), for every line $L = \{x, y, z\}$ of \mathcal{S}' the triple $\{f_x, f_y, f_z\}$ is a \mathcal{V} -line. Map every line of \mathcal{S}' to this corresponding line of $\mathcal{V}_{A,B}$. Since \mathcal{S}' and $\mathcal{V}_{A,B}$ have the same number of points and the same order $(2, 4)$, the above maps between the point and line sets of \mathcal{S}' and $\mathcal{V}_{A,B}$ are bijections and it defines an isomorphism between the two geometries by Theorem 2.3.2. Thus, every near polygon of order $(2, 4)$ that contains an isometrically embedded generalized hexagon \mathcal{S} isomorphic to $H(2)^D$ must be isomorphic to $\mathcal{V}_{A,B}$, which proves Theorem 6.1.4.

6.3.2. $G_2(4)$ Near Octagon

For this section let \mathcal{S}' be a near octagon with three points on each line containing a suboctagon \mathcal{S} isomorphic to HJ isometrically embedded in it. The valuation geometry \mathcal{V} of $\mathcal{S} \cong HJ$ is given in Tables 2.7 and 2.8. The main purpose of this section is to show that if \mathcal{S}' has order $(2, 10)$, then \mathcal{S} is isomorphic to the $G_2(4)$ near octagon constructed in Chapter 5. We use the same approach as in Section 6.3.1 of showing that if \mathcal{S}' has order $(2, 10)$, then it consists of points of type A , B or C and lines of type AAA, ABB, ACC, BBC or CCC , with each type occurring exactly once, which proves that up-to isomorphism there is at most one near octagon of order $(2, 10)$ containing HJ as an isometrically embedded subgeometry. Since in Chapter 5, we showed that HJ is indeed an isometrically embedded subgeometry of the $G_2(4)$ near octagon (see Lemma 5.4.9), this will prove the result. First we derive some general results that are true for any near octagon \mathcal{S}' with three points on each line that contains \mathcal{S} as a full isometrically embedded subgeometry and later restrict ourselves to the case when \mathcal{S}' has order $(2, 10)$. The following four lemmas are proved using the computer model of the valuation geometry of HJ constructed using the methods of Chapter 3.

Lemma 6.3.14. *Let f be a valuation of type C and let $g \neq f$ and $h \neq f$ be valuations of type B or C lying on distinct \mathcal{V} -lines through f . Then g and h are non-collinear.*

Lemma 6.3.15. *Let f be a valuation of type B and let $x \in \mathcal{S}$ be the unique point in \mathcal{O}_f . Then the map $\{f, g, h\} \mapsto \mathcal{O}_f \cup \mathcal{O}_g \cup \mathcal{O}_h$ is a bijection between the set of five \mathcal{V} -lines of type BBB through f and the set of five lines of \mathcal{H} through x .*

Lemma 6.3.16. *Let f be a valuation of type C . Then there is a unique \mathcal{V} -line $\{f, g, h\}$ of type CCC through f such that $\mathcal{O}_f \cup \mathcal{O}_g \cup \mathcal{O}_h$ is a line of HJ . For every other \mathcal{V} -line $\{f, g', h'\}$ of type CCC through f , the set $\mathcal{O}_f \cup \mathcal{O}_{g'} \cup \mathcal{O}_{h'}$ is a set of three pairwise non-collinear points.*

A \mathcal{V} -line $\{f, g, h\}$ of type CCC will be called *special* if $\mathcal{O}_f \cup \mathcal{O}_g \cup \mathcal{O}_h$ is a line of HJ . If that is not the case, then $\{f, g, h\}$ will be called an *ordinary* \mathcal{V} -line. This concept of *special* and *ordinary* is then extended to the lines of \mathcal{S}' that induce \mathcal{V} -lines of type CCC .

Lemma 6.3.17. *The subgeometry of \mathcal{V} defined on the type C valuations by the lines of type ACC and the ordinary lines of type CCC is connected.*

Remark. Lemmas 6.3.14 and 6.3.16 could alternatively be verified by a geometric reasoning inside the $G_2(4)$ near octagon, keeping in mind the construction of this near octagon using the valuation geometry of HJ (see Proposition 5.4.15).

Lemma 6.3.18. *Every point of \mathcal{S}' is at distance at most 2 from \mathcal{S} . Points of \mathcal{S} are of type A , points at distance 1 from \mathcal{S} are of type B or C and those at distance 2 are of type D or E .*

Proof. Since \mathcal{S} is isometrically embedded in \mathcal{S}' , all points of \mathcal{S} induce type A valuations. From Lemma 2.2.4 and the column M_f of Table 2.7 we see that the points in $\Gamma_1(\mathcal{S})$ cannot be of type A , but the points in $\Gamma_2(\mathcal{S})$ must be of type D or E . If $x \in \Gamma_1(\mathcal{S})$, then there exists a line through x that intersects \mathcal{S} , which must necessarily be of type ABB or ACC by Table 2.8, implying that x has type B or C . \square

Corollary 6.3.19. *Every point of \mathcal{S}' at distance 1 from \mathcal{S} is collinear with a unique point of \mathcal{S} .*

Proof. Such points are of type B or C and valuations of type B and C have exactly one point of value 0. \square

Lemma 6.3.20. *There are no points of type E in \mathcal{S}' .*

Proof. Let x be a type E point of \mathcal{S}' . By Lemma 6.3.18, x must be at distance 2 from \mathcal{S} . Let y be a neighbor of x which lies at distance 1 from \mathcal{S} . Then x has type B or C by Lemma 6.3.18. Since the valuations f_x and f_y are not equal, the line xy gives rise to a \mathcal{V} -line in the valuation geometry of HJ (see Theorem 2.3.2). But, by Table 2.8 there are no \mathcal{V} -lines with both type E and type T points on it, for $T \in \{B, C\}$. \square

Let x be a point of \mathcal{S}' at distance 1 from \mathcal{S} which by Lemma 6.3.18 is of type B or C . We will call the unique point of \mathcal{S} collinear with x (see Corollary 6.3.19) the *projection* of x , and denote it by $\pi(x)$. From now onward we implicitly use the fact that points at distance 1 from \mathcal{S} are of type B or C . For a line $L = \{x, y, z\}$ contained in $\Gamma_1(\mathcal{S})$ we define the projection $\pi(L)$ of L to be the set $\{\pi(x), \pi(y), \pi(z)\}$ of points of \mathcal{S} . Since \mathcal{S}' is a near polygon, $\pi(L)$ and L have the same size for every line L in $\Gamma_1(\mathcal{S})$. But, this projection may or may not be a line of \mathcal{S} .

Lemma 6.3.21. *Let x be a type B point of \mathcal{S}' and let y be a point on a line through x which does not intersect \mathcal{S} . Then y is at distance 1 from \mathcal{S} and the projections $\pi(x)$ and $\pi(y)$ are collinear.*

Proof. The point y must be of type B or C since there are no \mathcal{V} -lines containing both type B and type T points for $T \in \{D, E\}$ (see Table 2.8) and hence at distance 1 from \mathcal{S} by Lemma 6.3.18. The projections $\pi(x)$ and $\pi(y)$ have f_x -values 0 and 1, respectively. Since f_x is of type B , there are exactly ten points of \mathcal{S} that have f_x -value 1 (see Table 2.7). Clearly, every point in \mathcal{S} at distance 1 from $\pi(x)$ has f_x -value 1. Since \mathcal{S} has order $(2, 4)$, there are precisely ten such points and hence $\pi(y)$ must be one of them. \square

Corollary 6.3.22. *If x is a point of type B in \mathcal{S}' , then every line through x that does not intersect \mathcal{S} is parallel to and at distance 1 from a unique line of \mathcal{S} .*

Lemma 6.3.23. *Every type B point x of \mathcal{S}' is incident with a line of type BBC .*

Proof. Let x be a point of type B . Then every point of \mathcal{S} at distance 1 from $\pi(x)$ has f_x -value 1. Since $\mathcal{O}_{f_x} = \{\pi(x)\}$ every point of \mathcal{S} at distance 2 from $\pi(x)$ should have f_x -value 2, and since \mathcal{S} is a regular near octagon with parameters $(2, 4; 0, 3)$, there are 80 such points. By Table 2.7 there are 112 points of \mathcal{S} with f_x -value 2. Let y be one of the other $112 - 80 = 32$ points with f_x -value 2 at distance at least 3 from $\pi(x)$. Since $f_x(y) = 2$, we have $d(x, y) = 3$. Let x, u, v, y be a path of length 3 connecting x and y . By Lemmas 6.3.18 and 6.3.21, the point u has type A , B or C . We will show that u is of type C , hence proving that the line xu is of type BBC .

If u is of type A , then $u = \pi(x)$, which would be in contradiction with $d(\pi(x), y) > 2$. Suppose u is of type B , and hence at distance 1 from \mathcal{S} . From Corollary 6.3.22 we see that $\pi(u)$ and $\pi(x)$ are collinear (or equal). We cannot have $v = \pi(u)$ as that would imply that $d(\pi(x), y) \leq 2$. Therefore, v lies outside \mathcal{S} and y must be equal to $\pi(v)$. Again by Corollary 6.3.22 $y = \pi(v)$ and $\pi(u)$ must be collinear (or equal), which contradicts the fact that $d(\pi(x), y) > 2$. So, u is of type C . \square

Corollary 6.3.24. *There exist type C points in \mathcal{S}' .*

Proof. As there exist points at distance 1 from \mathcal{S} , there exist points of type B or C . The existence of type B points implies the existence of type C points by Lemma 6.3.23. \square

Lemma 6.3.25. *Let x be a point of \mathcal{S}' of type C and let L_1, L_2 be two distinct lines of type CCC through x . Then the \mathcal{V} -lines corresponding to L_1 and L_2 must be distinct.*

Proof. Let $L_1 = \{x, y, z\}$ and $L_2 = \{x, y', z'\}$. Assume that they correspond to the same \mathcal{V} -line so that $f_y = f_{y'}$ and $f_z = f_{z'}$. Let $u := \pi(z) = \pi(z')$. Since $xzuz'$ is a quadrangle, the point y must be collinear with the third point on the line uz' , call it v . Therefore, the valuations f_v and $f_y = f_{y'}$ are collinear in \mathcal{V} . The collinearity of f_v and $f_{y'}$ in \mathcal{V} contradicts Lemma 6.3.14 by taking $f = f_{z'}$, $g = f_{y'}$ and $h = f_v$. \square

On the valuations of type C we can define a subgeometry of \mathcal{V} induced by the lines of type ACC and the ordinary lines of type CCC . Let this subgeometry be denoted by \mathcal{V}_C and its collinearity graph by Γ_1 . Similarly, we can define a subgeometry \mathcal{S}'_C of \mathcal{S}' by taking the points of type C and the lines that correspond to lines of \mathcal{V}_C . Let Γ_2 be the collinearity graph of \mathcal{S}'_C . Since type C points exist in \mathcal{S}' , the graph Γ_2 is nonempty.

Lemma 6.3.26. *The graph Γ_2 is a cover of the graph Γ_1 by the map $x \mapsto f_x$.*

Proof. Let x be a point of type C in \mathcal{S}' . There are 16 points of f_x -value 1 that are not collinear with $\pi(x)$ (see Table 2.7). Denote this set of 16 points by U . If $u \in U$, then $d(x, u) = 2$. Denote by v a common neighbor of x and u . Since $v \neq \pi(x)$, $v \notin \mathcal{S}$ and $u = \pi(v)$. Since $u \neq \pi(x)$, the point v cannot be contained on the line $x\pi(x)$, and so xv is a line disjoint from \mathcal{S} . Since $\pi(x)$ and $\pi(v) = u$ are not collinear, Lemma 6.3.21 implies that v has type C and hence that xv is a (necessarily ordinary) line of type CCC .

By Lemmas 6.3.16, 6.3.25 and Table 2.8 there are at most 8 ordinary lines of type CCC through x , each of which determines two points of the set U . Since $|U| = 16$, it follows that there are precisely 8 ordinary lines of type CCC through x and they correspond bijectively to the 8 ordinary \mathcal{V} -lines of type CCC through f_x . This proves that the map $x \mapsto f_x$ is a local isomorphism between Γ_2 and Γ_1 . The fact that this map is surjective now follows from the connectedness of Γ_1 , see Lemma 6.3.17. \square

Corollary 6.3.27. *If Γ_2 is an i -cover of Γ_1 for some $i \geq 1$, then each valuation of type C is induced by precisely i type C points of \mathcal{S}' . As a consequence, through each point of \mathcal{S} , there are precisely $5i$ lines of type ACC .*

Proof. By Table 2.8 there are precisely 5 \mathcal{V} -lines of type ACC through a given valuation of type A . Since each type C valuation occurs exactly i times, we have $5i$ type ACC lines in \mathcal{S}' through a given point of \mathcal{S} . \square

Remark. All results in this section so far are valid for a general near octagon of order $(2, t)$ that contains an isometrically embedded sub near octagon isomorphic to \mathbf{HJ} . In the following lemma, we need the fact that \mathcal{S}' has order $(2, 10)$.

Lemma 6.3.28. *If \mathcal{S}' is of order $(2, 10)$, then each valuation of type $T \in \{B, C\}$ is induced exactly once by a point of \mathcal{S}' .*

Proof. Let \mathcal{S}' be of order $(2, 10)$ and let x be an arbitrary point of \mathcal{S} . Then there are exactly $11 - 5 = 6$ lines through x that are not contained in \mathcal{S} , each of which has type ACC or ABB . Since type C points exist by Corollary 6.3.24, it follows from Corollary 6.3.27 that there are precisely 5 lines of type ACC through x in \mathcal{S}' , and hence the graph Γ_2 is a 1-cover of Γ_1 . Now the 6-th line through x which is not contained in \mathcal{S} must be of type ABB . Therefore, through every point of \mathcal{S} there are 5 lines of type ACC and a unique line of type ABB . This shows that for every valuation f of type B , we can find the unique point of \mathcal{S}' that induces f by first getting the point y of \mathcal{S} that induces the type A valuation on the unique \mathcal{V} -line of type ABB through f (see Table 2.8), and then picking the point on the unique line of type ABB through x in \mathcal{S}' that induces the valuation f . \square

For the rest of this section assume that \mathcal{S}' has order $(2, 10)$. From Lemma 6.3.28 we know that both type B and type C points exist in \mathcal{S}' and each type B or type C valuation of \mathcal{S} is induced by a unique point of \mathcal{S}' . Let x be a point of type B in \mathcal{S}' and let L_x be the unique line joining x and $\pi(x)$. From Corollary 6.3.22 it follows that every other line through x gives rise to a quad in \mathcal{S}' that intersects \mathcal{S} and contains L_x .

Lemma 6.3.29. *Let Q be a quad of \mathcal{S}' that intersects \mathcal{S} nontrivially. Then $Q \cap \mathcal{S}$ is either a singleton or a line.*

Proof. The proof is similar to that of Lemma 6.3.8. \square

Lemma 6.3.30. *Let Q be a quad of \mathcal{S}' that is not a grid and that intersects \mathcal{S} in a line L . Then there must exist points of type B and points of type C in $Q \setminus L$.*

Proof. For the sake of contradiction assume that all points of $Q \setminus L$ are of a fixed type $T \in \{B, C\}$. Let x be a point of L . Since Q is not a grid, there exist two lines $L_1 = \{x, y, z\}$ and $L_2 = \{x, y', z'\}$ through x with $y, y', z, z' \in Q \setminus L$. Let w be a common neighbor of z and z' in Q which is different from x . Then $w \in Q \setminus L$. From Lemma 6.3.28 it follows that the lines wz and wz' correspond to distinct \mathcal{V} -lines, which are of type TTT by our assumption. Also note that $\pi(wz) = \pi(wz') = L$. This contradicts Lemma 6.3.15 for $T = B$ and Lemma 6.3.16 for $T = C$. \square

Lemma 6.3.31. *There are no $Q(5, 2)$ -quads in \mathcal{S}' that meet \mathcal{S} in a line.*

Proof. Let Q be a $Q(5, 2)$ -quad that meets \mathcal{S} in a line L . By Lemma 6.3.30 there is a point x of type C in Q . There is a unique line through x that intersects \mathcal{S} in L , and hence lies in Q . Every other line through x which is contained in Q projects to L . By Lemmas

6.3.16 and 6.3.25 there is at most one line of type CCC through x in Q . From Table 2.8 and Lemma 6.3.28 it follows that there is at most one line of type BBC through x in Q . Therefore, in total we have at most three lines through x in Q which contradicts the fact that the order of a $Q(5, 2)$ -quad is $(2, 4)$. \square

Lemma 6.3.32. *Let x be a point of type B in \mathcal{S}' . Then x cannot be contained in two lines L_1, L_2 such that L_1 has type BBB , L_2 has type BBC and $\pi(L_1) = \pi(L_2)$.*

Proof. Let $L_1 = \{x, y, z\}$ and $L_2 = \{x, y', z'\}$ be two such lines, such that $\pi(y) = \pi(y')$ and $\pi(z) = \pi(z')$. Say L_1 is of type BBC and L_2 of type BBB . Without loss of generality assume that z is of type C . Then the lines $y\pi(y)$ and $y'\pi(y')$ are of type ABB . By Table 2.8 there is only one \mathcal{V} -line of type ABB through a valuation of type A , and hence f_y is equal to $f_{y'}$ or $f_{y''}$ where y'' is the third point (of type B) on the line $y'\pi(y')$. This contradicts Lemma 6.3.28. \square

Lemma 6.3.33. *Let x be a point of type B in \mathcal{S}' . Then*

- (1) *x is incident with a unique line of type ABB and ten lines of type BBC ;*
- (2) *these ten type BBC lines through x correspond bijectively to the ten BBC lines of the valuation geometry \mathcal{V} through f_x , and they are partitioned into pairs by five $W(2)$ -quads passing through the line of type ABB through x .*

Proof. There is a unique line through x that intersects \mathcal{S} , namely the line joining x and $\pi(x)$. Every other line through x is of type BBB or BBC which is entirely contained in $\Gamma_1(\mathcal{S})$ and is parallel to a line through $\pi(x)$ in \mathcal{S} (see Corollary 6.3.22). Let S denote the set of these other lines through x . By Lemma 6.3.28, distinct lines in S correspond to distinct \mathcal{V} -lines. Let there be i lines of type BBB in S , with $i \leq 5$ by Table 2.8. Since we cannot have two lines of type BBC and BBB in S projecting to the same line of \mathcal{S} by Lemma 6.3.32 and since there are no $Q(5, 2)$ -quads by Lemma 6.3.31, there are at most $2(5 - i)$ lines of type BBC in S , and hence in total at most $2(5 - i) + i + 1 = 11 - i$ lines through x . Therefore, we have $i = 0$ and each of the 5 lines of \mathcal{S} through $\pi(x)$ is parallel to exactly 2 lines of S . This gives rise to 5 $W(2)$ -quads through the line $x\pi(x)$, that partition S into pairs. \square

We are now ready to prove Theorem 6.1.5. From Lemma 6.3.33 it follows that there are no lines of type BBB in \mathcal{S}' . Since each of the type A , B and C valuations is induced by a unique point of \mathcal{S}' and each \mathcal{V} -line of type AAA , ABB and ACC is induced by a unique line of \mathcal{S}' , it suffices to show that also every \mathcal{V} -line of type BBC and CCC is induced by a unique line of \mathcal{S}' , and that type D points do not exist in \mathcal{S}' (we have already proved in Lemma 6.3.20 that type E points do not exist).

Let $\{f, g, h\}$ be a \mathcal{V} -line of type BBC where f is of type B . Let x be the unique point in \mathcal{S}' with $f_x = f$. By Lemma 6.3.33, there exists a line $L = \{x, y, z\}$ such that $f_y = g$ and $f_z = h$. This shows that each \mathcal{V} -line of type BBC is induced by a necessarily unique line of \mathcal{S}' .

Now, let x be a point of type C . Since Γ_2 is a 1-cover of Γ_1 , there exist eight ordinary lines of type CCC through x that bijectively correspond to the eight ordinary \mathcal{V} -lines of type CCC through f_x . By Table 2.8, there exists a unique \mathcal{V} -line of type BBC through f_x , implying that in \mathcal{S}' there is a unique line L of type BBC through x . By Lemma 6.3.33 L lies in a $W(2)$ -quad Q , which must also contain the unique line of type ACC through x . The third line in Q through x must be a special line of type CCC as there is a unique

type BBC line through x and none of the ordinary type CCC lines through x projects to a line of \mathcal{S} . Therefore, the unique special \mathcal{V} -line of type CCC through f_x is induced by a line of \mathcal{S}' . Since we have accounted for all 11 lines through a point of type C , there cannot be any lines of type CDD , and hence there cannot be any points of type D in \mathcal{S}' . This completes the proof as we have shown that \mathcal{S}' is isomorphic to the $G_2(4)$ near octagon.

Part II.

Polynomial Method

7. Introduction

The *polynomial method* is an umbrella term for different techniques involving polynomials which have been used to solve several problems in finite geometry, discrete geometry, extremal combinatorics and additive number theory. One of the general philosophies behind this method is to associate a set of polynomials (possibly a single polynomial), to a combinatorial object that we want to study, and then use some properties of these polynomials to describe the combinatorial object. For a concrete example, let us go through Koornwinder’s proof of the absolute bound on equiangular lines.

A set of lines in the Euclidean space \mathbb{R}^n through the origin (or any other fixed point) is called equiangular if the angle between every pair of distinct lines in the set is the same. For example, joining the opposite vertices of a regular hexagon in the plane \mathbb{R}^2 , we get three equiangular lines.

At most how many equiangular lines can there be in \mathbb{R}^n ?

This question was addressed by Gerzon (as reported by Lemmens and Seidel in [102]), who proved that there are at most $\binom{n+1}{2}$ equiangular lines in \mathbb{R}^n . Thus in particular, the regular hexagon example gives us the maximum possible equiangular lines in \mathbb{R}^2 . But in general this bound is not sharp. In 1976, Koornwinder gave an “almost trivial proof” [100] of Gerzon’s bound by giving a bijective correspondence between the set of equiangular lines in \mathbb{R}^n and a linearly independent set of polynomials lying in an $\binom{n+1}{2}$ dimensional vector space.¹ This correspondence is as follows. Let L_1, \dots, L_k be k equiangular lines in \mathbb{R}^n and let u_1, \dots, u_k be unit vectors on these lines, chosen arbitrarily. Then we have $\langle u_i, u_j \rangle^2 = \alpha$ for $i \neq j$ where α is a fixed real number in the interval $[0, 1)$. For $i \in \{1, \dots, k\}$, define $f_i \in \mathbb{R}[t_1, \dots, t_n]$ by $f_i(t_1, \dots, t_n) = (\langle u_i, (t_1, \dots, t_n) \rangle)^2 - \alpha^2(t_1^2 + \dots + t_n^2)$. Now since $f_i(u_j) = 0$ for $i \neq j$ and $f_i(u_i) = 1 - \alpha^2 \neq 0$ for all i , it is easy to see that f_1, \dots, f_k are linearly independent polynomials in the vector space $\mathbb{R}[t_1, \dots, t_n]$. As these polynomials lie in the $\binom{n+1}{2}$ dimensional subspace of $\mathbb{R}[t_1, \dots, t_n]$ spanned by the monomial set $\{t_i t_j\}_{1 \leq i < j \leq n}$, we get $k \leq \binom{n+1}{2}$.

The argument above can also be seen as an example of the linear algebra method in combinatorics, which has been discussed in much detail in the influential (unfinished) manuscript of Babai and Frankl [8], and more recently in the beautiful book by Matoušek [108].

Another important way of using polynomials is to capture the combinatorial object via zeros of polynomials (or in general, algebraic varieties). One of the earliest examples here is the determination of the minimum size of an affine blocking set by Jamison in 1977 [93]. The problem is to find the minimum number of points required to “block” every hyperplane of the affine space \mathbb{F}_q^n . Clearly $n(q - 1) + 1$ points suffice (by taking

¹In fact, this is one of the earliest papers that played an important role in the development of polynomial method.

all points that lie on one of the n axes), but can we do better? Jamison proved that we cannot, and his polynomial method proof can be sketched as follows: (a) first consider the dual problem, which is equivalent to finding the minimum number of hyperplanes required to cover all points of \mathbb{F}_q^n except the origin, (b) then identify \mathbb{F}_q^n with the finite field \mathbb{F}_{q^n} , (c) finally associate each hyperplane with the minimal polynomial over \mathbb{F}_{q^n} that vanishes on the hyperplane to show (using the theory of linearized polynomials [103, Chapter 3]) that if the number of hyperplanes is less than $n(q-1)+1$, then the polynomial $t^{q^n-1} - 1 = \prod_{\alpha \in \mathbb{F}_{q^n}^\times} (t - \alpha)$, does not divide the product of these polynomials corresponding to the hyperplanes. This technique of using polynomials over finite fields to solve finite geometrical problems came to be known as the “Jamison method” and it saw several applications (see for example, the surveys [25] and [39]).

Brouwer and Schrijver gave another proof of Jamison’s theorem in [37] where they also started by considering the dual problem of hyperplane coverings but then proceeded by a much simpler argument involving multivariate polynomials over finite fields. Their approach was in fact quite similar to Chevalley’s proof of the famous Chevalley-Warning theorem [47] using reduced polynomials. We will see in Chapters 8 and 9 how both of these results are linked together by the notion of grid reduction, and in particular by the Lemma that a polynomial $\mathbb{F}_q[t_1, \dots, t_n]$ which vanishes on all points of \mathbb{F}_q^n except one must have degree at least $n(q-1)$.² The Chevalley-Warning theorem, which is a statement on the zero set of a collection of polynomials over a finite field, has also found several applications in combinatorics. Alon, Friedland and Kalai used it to prove that every 4-regular graph plus an edge contains a 3-regular subgraph [2]. Later, Bailey and Richter [9] used the Chevalley-Warning theorem to give a new proof of the famous Erdős-Ginzburg-Ziv theorem in additive number theory [79]. Recently, Clark, Forrow and Schmitt [51] have shown that the Chevalley-Warning theorem and its combinatorial applications can be derived, and further generalized, using a result of Alon and Füredi from 1993 [3, Theorem 5]. We will devote Chapter 9 to this Alon-Füredi Theorem, where we generalize the result and give a new simple proof. We also show how this result is linked to several other topics like coding theory, finite geometry and polynomial identity testing.

An important tool in the polynomial method involving zeros of polynomials is a result called *Combinatorial Nullstellensatz* [1], which was developed by Alon and his collaborators [4, 5]. This powerful tool and its generalizations have been used extensively to solve several problems in additive number theory (see [134, Chapter 9] for a survey) and more recently in some other areas as well [95, 96]. In [49], Clark revisited Alon’s Combinatorial Nullstellensatz and showed how its proof can be seen as a “restricted variable analogue” of Chevalley’s proof of the Chevalley-Warning Theorem. He further generalized this result to commutative rings (adding certain extra conditions) and made it clear how many of the ideas involved are related to the notion of grid reduction. In [13], Ball and Serra introduced a related result which they called *Punctured Combinatorial Nullstellensatz*. This result was proved using Alon’s Combinatorial Nullstellensatz, and it has several combinatorial applications of its own. We will give another proof of this result in Chapter 8 by directly using the notion of grid reduction, and then use this result to prove a new generalization of the Chevalley-Warning theorem which we call the *Punctured Chevalley-Warning Theorem*. In fact, this result generalizes Brink’s Restricted Variable Chevalley-Warning theorem [29].

²Note that we can find a polynomial of degree q which vanishes on all points, and therefore this is a big jump!

In recent years, there has been a lot of interest in the polynomial method as a result of Dvir's two-page proof of the finite field Kakeya problem [76]³ which involved an easy polynomial argument, and the developments that followed. Many experts worked on the finite field Kakeya problem using different techniques involving algebraic geometry and Fourier analysis, but made only partial progress towards a solution to this problem. And thus it was a great surprise to the mathematical community that such an easy polynomial argument could resolve this famous open problem. Ideas originating from Dvir's work have lead to several important advancements in mathematics, including the big breakthrough in the famous Erdős distinct distances problem by Guth and Katz [87]. It is interesting to note that Dvir's polynomial technique is quite different from the techniques we have mentioned so far in this introduction as it involved polynomial interpolation instead of constructing explicit polynomials. For more details on this, we recommend the surveys by Dvir [77] and Tao [133], and the recent book by Guth [86]. Another example where a combinatorial problem is solved using polynomial interpolation, combined with a geometric argument, is Segre's classical theorem on ovals in finite projective planes [123]. Interestingly, the so-called "lemma of tangents" from [123] was used in combination with the Jamison/Brouwer-Schrijver bound on affine blocking sets by Blokhuis and Korchmáros [28] to solve the Kakeya problem in 2 dimensions. Segre's result (and his lemma of tangents) has been generalized further to higher dimensional finite projective spaces by Ball [10]. For more on polynomial method in finite geometry, see the survey by Ball [11].

Notation

We will be using the following notation in Chapters 8 and 9. For us rings will always be commutative with a multiplicative identity denoted by 1. R denotes an arbitrary ring, and F an arbitrary field. For an n -variable polynomial $f \in R[t_1, \dots, t_n]$ we let $\deg f$ denote the total degree of f and $\deg_{t_i} f$ the degree of f when it is treated as an element of $R'[t_i]$ where $R' = R[t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n]$. If f is the zero polynomial of $R[t_1, \dots, t_n]$, then we will assume that $\deg f = -\infty$, i.e., the zero polynomial is assumed to have a degree smaller than every non-zero polynomial. We will sometimes use the shorthand $R[t]$ for $R[t_1, \dots, t_n]$ when n is clear from the context.

A *finite grid* in R^n (or F^n) is a subset A of the form $A = A_1 \times \dots \times A_n$ where A_1, \dots, A_n are finite nonempty subsets of R (or F). For $f \in R[t_1, \dots, t_n]$ and $A \subseteq R^n$, we put

$$Z_A(f) = \{x \in A \mid f(x) = 0\} \text{ and } \mathcal{U}_A(f) = \{x \in A \mid f(x) \neq 0\}.$$

For $N, a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{Z}^+$, with $N \leq \sum_{i=1}^n a_i$ and $1 \leq b_i \leq a_i$ for all i , we denote by $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N)$ a certain combinatorial quantity related to (restricted) distribution of N balls in n bins, which will be defined in Section 9.2. When $b_1 = \dots = b_n = 1$ we denote this quantity by $\mathbf{m}(a_1, \dots, a_n; N)$.

³The problem asks if there exists a constant c_n such that the minimum number of points in \mathbb{F}_q^n that contain a line in every direction is at least $c_n q^n$.

8. Grid Reduction

8.1. Basics

One of the most basic results on polynomials is the fact that a single variable polynomial of degree d over an integral domain has at most d zeros. This result is clearly not true over arbitrary rings as one can see from the example $f = t^2 - 1 \in (\mathbb{Z}/8\mathbb{Z})[t]$ which has 4 zeros in $\mathbb{Z}/8\mathbb{Z}$. To rectify this situation, we restrict to certain subsets of R that we now define. Following [49] and [121], we say that a subset S of R satisfies *Condition (D)* if for every pair of distinct elements α, β in S , the element $\alpha - \beta$ of R is not a zero divisor. We say that a finite grid $A = A_1 \times \cdots \times A_n \subseteq R^n$ satisfies Condition (D) if every A_i satisfies Condition (D). Note that when R is an integral domain, this condition is automatically satisfied and in fact this is the case which most commonly occurs in combinatorial applications (with A being the binary grid $\{0, 1\}^n$). We will see that any single variable polynomial of degree d over a ring R has at most d zeros in any subset of R that satisfies Condition (D). But first, we mention some elementary results on polynomials over rings whose proofs are left to the reader.

Lemma 8.1.1. *Let $f, g, h \in R[t_1, \dots, t_n]$ with $f = g + h$.*

- (1) *If $\deg f \geq \deg g$, then $\deg f \geq \deg h$.*
- (2) *For all $i \in \{1, \dots, n\}$, if $\deg_{t_i} f \geq \deg_{t_i} g$, then $\deg_{t_i} f \geq \deg_{t_i} h$.*

Lemma 8.1.2. *Let $f, g, h \in R[t_1, \dots, t_n]$ with g monic¹ and $f = gh$. Then $\deg f = \deg g + \deg h$.*

Lemma 8.1.3 (Euclidean Division Algorithm). *Let $f, g \in R[t]$ be single variable polynomials with g monic and nonzero. Then there exist polynomials $h, r \in R[t]$ such that $f = gh + r$ and $\deg r < \deg g$.*

Corollary 8.1.4. *Let $f \in R[t]$ be a single variable nonzero polynomial. If $S \subseteq R$ is such that S satisfies Condition (D), then f has at most $\deg f$ zeros in S .*

Proof. Let $\alpha \in S$ be a zero of f . We can write $f = (t - \alpha)h + r$ with $\deg r < 1$ (so $r \in R$). Substituting α in the equation we see that $r = 0$. Since $f \neq 0$, we have $h \neq 0$. Let $\beta \in S \setminus \{\alpha\}$. Then $f(\beta) = (\beta - \alpha)h(\beta)$. Since $\beta - \alpha$ is not a zero divisor, $f(\beta) = 0$ if and only if $h(\beta) = 0$. Moreover, $\deg h = \deg f - 1$ by Lemma 8.1.2. Thus, the result follows by induction on $\deg f$. \square

¹We call a multivariate polynomial monic if the coefficient of its leading term, with respect to some monomial ordering, is 1. We could have also assumed that $g \in R[t_i]$ for some $i \in \{1, \dots, n\}$ as that's the only case which will be used later in the thesis.

In other words, if f has more than $\deg f$ zeros in a set $S \subseteq R$ satisfying Condition (D), then $f = 0$. The following lemma is a generalization of this fact to multivariate polynomials.

Lemma 8.1.5. *Let R be a ring and let A_1, \dots, A_n be nonempty finite subsets of R that satisfy Condition (D). Put $A := A_1 \times \dots \times A_n$. Let $f \in R[t_1, \dots, t_n]$ such that for all $i \in \{1, \dots, n\}$ we have $\deg_{t_i} f \leq |A_i| - 1$. If $f(x) = 0$ for all $x \in A$, then $f = 0$.*

Proof. We prove this by induction on n . The base case is clear from Corollary 8.1.4. For the inductive step write

$$f = \sum_{i=0}^{d_n} f_i(t_1, \dots, t_{n-1}) t_n^i.$$

where $d_n = \deg_{t_n} f$. Let $x' \in A' = A_1 \times \dots \times A_{n-1}$. Since $f(x) = 0$ for all A we see that the single variable polynomial $f(x', t_n)$ vanishes everywhere on A_n . Since $\deg f(x', t_n) = d_n < |A_n|$, and A_n satisfied Condition (D), Corollary 8.1.4 implies that $f(x', t_n)$ is the zero polynomial. This implies that for all i , $f_i(x') = 0$ for all $x' \in A'$. By the induction hypothesis, $f_i = 0$ for all i and hence $f = 0$. \square

Definition. Let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n . Then a polynomial $f \in R[t_1, \dots, t_n]$ is called *A-reduced* if for all $i \in \{1, \dots, n\}$ we have $\deg_{t_i} f \leq |A_i| - 1$. We denote the set of all *A-reduced* polynomials by \mathcal{R}_A .

Therefore, Lemma 8.1.5 says that if a grid A in R^n satisfies Condition (D), then every *A-reduced* polynomial which vanishes at all points of A must be equal to the zero polynomial. The following Lemma is a multivariate generalization of Lemma 8.1.3 and it will help us define the notion of *grid reduction*.

Lemma 8.1.6. *Let $f \in R[t_1, \dots, t_n]$ and let g be a monic nonzero polynomial in $R[t_i]$ for some $i \in \{1, \dots, n\}$. Then, there exist $h, r \in R[t_1, \dots, t_n]$ such that*

$$f = gh + r$$

and $\deg_{t_i} r < \deg g$. Moreover, $\deg h \leq \deg f - \deg g$, and $\deg r \leq \deg f$.

Proof. Without loss of generality, take $i = n$. Let $R' = R[t_1, \dots, t_{n-1}]$. Then f and g can be seen as polynomials in $R'[t_n]$. If $\deg_{t_n} r < \deg g$, then we can take $h = 0$ and $r = f$. So, assume that $\deg_{t_n} r \geq \deg g$. By Lemma 8.1.3 $f = gh + r$ for some polynomials $h, r \in R'[t_n]$ such that $\deg r < \deg g$ and $h \neq 0$. If we now see h and r as polynomials in $R[t_1, \dots, t_n]$ we get that $\deg_{t_n} r < \deg g$. Since $\deg_{t_n} r < d = \deg g$, for every monomial $t_1^{e_1} \dots t_n^{e_n}$ in h with $e_1 + \dots + e_n = \deg h$ there exists a monomial $t_1^{e_1} t_2^{e_2} \dots t_n^{e_n + d}$ in f . This shows that $\deg f \geq \deg h + \deg g$. Now from Lemmas 8.1.1 and 8.1.2 it follows that $\deg f \geq \deg r$. \square

Let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n . For each $i \in \{1, \dots, n\}$, let $\varphi_i \in R[t_i]$ be the polynomial $\prod_{\lambda \in A_i} (t_i - \lambda)$ of degree $|A_i|$. We let Φ denote the ideal of $R[t_1, \dots, t_n]$ generated by the polynomials $\varphi_1, \dots, \varphi_n$. Then clearly $\Phi \subseteq I(A)$, where $I(A)$ is the ideal of all polynomials in $R[t_1, \dots, t_n]$ that vanish on A . By repeated applications of Lemma 8.1.6, for every polynomial $f \in R[t_1, \dots, t_n]$ we can find polynomials $g_1, \dots, g_n \in R[t_1, \dots, t_n]$ and $r \in \mathcal{R}_A$ such that

$$(1) \quad f = g_1 \varphi_1 + \dots + g_n \varphi_n + r;$$

- (2) $\deg f \geq \deg r$;
 (3) $\deg g_i \leq \deg f - \deg \varphi_i$ for all $i \in \{1, \dots, n\}$.

This process is what we call the *grid reduction* of the polynomial f with respect to the finite grid A . One can also define a more general notion of “Cartesian Reduction” by taking φ_i ’s to be arbitrary monic polynomials in single variables, as was done by Clark in [49, Section 3]. It is also easy to show that the “remainder” r that one gets by this process of grid reduction is unique by proving that the ideal Φ does not contain any A -reduced polynomials (see [49, Proposition 10]). We will denote this unique A -reduced polynomial that one gets after grid-reduction of f by $r_A(f)$. From Lemma 8.1.5 we have the following result, which Clark calls the *CATS Lemma* [49, Theorem 12] (after Chevalley [47], Alon-Tarsi [5] and Schauz [121]).

Lemma 8.1.7. *Let R be a ring and let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n . Then the following are equivalent:*

- (1) A satisfies Condition (D);
 (2) If $r \in \mathcal{R}_A$ and $r(x) = 0$ for all $x \in A$, then $r = 0$;
 (3) $\Phi = I(A)$.

In particular, if A satisfies Condition (D), then for every polynomial f , there exists a unique A -reduced polynomial $r_A(f)$ such that $f(x) = r_A(f)(x)$ for all $x \in A$. Moreover, $\deg f \geq \deg r_A(f)$.

Proof. (1) \implies (2) by Lemma 8.1.5. Let $f \in I(A)$. From grid reduction, we get $f = (\sum g_i \varphi_i) + r$, where r is A -reduced. Since $f(x) = 0$ for all $x \in A$, we see that $r(x) = 0$ for all $x \in A$, and hence $r = 0$ if (2) is assumed to be true. Therefore, (2) \implies (3). Now say (1) is false, and let $\alpha, \beta \in A_i$ for some i with $\alpha \neq \beta$, such that there exists a $\gamma \in R \setminus \{0\}$ with $\gamma(\alpha - \beta) = 0$. Then the polynomial $f = \gamma \prod_{\lambda \in A_i \setminus \{\alpha\}} (t_i - \lambda)$ is an A -reduced polynomial which is contained in $I(A)$. But no A -reduced polynomial can be contained in Φ , and hence $I(A) \neq \Phi$. Therefore, (3) \implies (1). \square

8.2. Applications of Grid Reduction

We first prove a lemma on zeros of polynomials in a finite grid, from which the Chevalley-Warning theorem, Jamison/Brouwer-Schrijver theorem on affine blocking sets, and the Alon-Füredi theorem on hyperplanes covering [3, Theorem 2] follow easily. We include the proof of the Chevalley-Warning theorem and discuss the other two results in Chapter 9, where they will follow from a more general result. Later on, we will see some other proofs of this Lemma as well.

Lemma 8.2.1 (All Except One). *Let $f \in R[t_1, \dots, t_n]$ and let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n that satisfies Condition (D). If f vanishes on all points of A except one, then $\deg f \geq \sum (|A_i| - 1)$.*

Proof. Let $a = (a_1, \dots, a_n)$ be the point where f doesn't vanish. The polynomial

$$g(\underline{t}) = \prod_{i=1}^n \prod_{\lambda \in A_i \setminus \{a_i\}} (t_i - \lambda)$$

of degree $\sum(|A_i| - 1)$ is an A -reduced polynomial that vanishes everywhere on A except at a . Therefore, the polynomial $g(a)r_A(f) - f(a)g$ is an A -reduced polynomial that vanishes everywhere on A , which implies that $g(a)r_A(f) = f(a)g$ by Lemma 8.1.7. Since $\deg r_A(f) = \deg g$ and $\deg f \geq \deg r_A(f)$, we get $\deg f \geq \deg g = \sum_{i=1}^n (|A_i| - 1)$. \square

Remark. When R is a field, Lemma 8.2.1 appears implicitly in [27] where several applications of this lemma in finite geometry are explored.

Theorem 8.2.2 (Chevalley-Warning Theorem [47, 139]). *Let $f_1, \dots, f_r \in \mathbb{F}_q[t_1, \dots, t_n]$ be such that $\sum_{i=1}^r \deg f_i < n$ and let Z be the set of common zeros of f_1, \dots, f_r in \mathbb{F}_q^n . Then $|Z| \neq 1$.*

Proof. Define a polynomial $f \in \mathbb{F}_q[t_1, \dots, t_n]$ by $f = \prod_{i=1}^r (1 - f_i^{q-1})$. Then since $\lambda^{q-1} = 1$ for every $\lambda \in \mathbb{F}_q \setminus \{0\}$, we have $f(x) = 1$ if $x \in Z$ and $f(x) = 0$ if $x \notin Z$. Say $|Z| = 1$. Then f vanishes on all points of the grid \mathbb{F}_q^n except one point. Therefore, by Lemma 8.2.1, we have $\deg f = (q-1) \sum_{i=1}^r \deg f_i \geq n(q-1)$. This contradicts the fact that $\sum_{i=1}^r \deg f_i < n$. \square

In fact, one can as easily prove Brink's generalization of Chevalley-Warning theorem [29] using Lemma 8.2.1 (also see [49, Theorem 19]).

Theorem 8.2.3 (Restricted Variable Chevalley-Warning Theorem). *Let $f_1, \dots, f_r \in \mathbb{F}_q[t_1, \dots, t_n]$ and let A_1, \dots, A_n be nonempty subsets of \mathbb{F}_q . Let $Z_A = \{x \in \prod A_i \mid \forall j, f_j(x) = 0\}$. If $(q-1) \sum \deg f_j < \sum (|A_i| - 1)$, then $|Z_A| \neq 1$.*

Proof. Again let $f = \prod (1 - f_j^{q-1})$ be the polynomial of degree $(q-1) \sum \deg f_j$ which is equal to 1 on all common zeros of f_j 's and 0 otherwise. Then Z_A corresponds to the set of non-zeros of f in $A = A_1 \times \dots \times A_n$. Say $|Z_A| = 1$. Then by Lemma 8.2.1 we have $\deg f = (q-1) \sum \deg f_j \geq \sum (|A_i| - 1)$, which is a contradiction. \square

Next we see a proof of Alon's Combinatorial Nullstellensatz [1] that uses the ideas of grid reduction developed so far. We note that this proof is essentially the same as the proof given by Alon in [1]. Its generalization over rings is due to Schauz [121] and the idea of looking at this proof in the following way is due to Clark [49].

Theorem 8.2.4 (Combinatorial Nullstellensatz for Rings). *Let $f \in R[t_1, \dots, t_n]$ and let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n that satisfies Condition (D). For all $i \in \{1, \dots, n\}$, define $\varphi_i = \prod_{\lambda \in A_i} (t_i - \lambda)$.*

- (1) *If f vanishes on all points of A , then there exist g_1, \dots, g_n such that $f = \sum g_i \varphi_i$ and $\deg g_i \leq \deg f - \deg \varphi_i$ for all i .*
- (2) *If there is a monomial term $\prod t_i^{d_i}$ in f with the property that $d_1 + \dots + d_n = \deg f$ and $d_i < \deg \varphi_i = |A_i|$ for all i , then f doesn't vanish everywhere on A .*

Proof. Reduce f modulo the grid A . Then $r_A(f)$ vanishes everywhere on A , and hence by Lemma 8.1.5 it must be zero. For (2) observe that such a monomial term is not affected by grid reduction, and hence the reduced form $r_A(f)$ is non-zero. If f vanishes

everywhere on A , then $r_A(f)$ also vanishes everywhere on A , which is not possible as it would contradict Lemma 8.1.5. ² \square

We now give a simple proof of a variation of Ball and Serra's Punctured Combinatorial Nullstellensatz [13, 14] over rings.³ This generalization is due to Clark [50], but our proof is different from his.

Lemma 8.2.5. *Let $A = A_1 \times \cdots \times A_n$ be a finite grid in R^n that satisfies Condition (D). Let $f \in R[t_1, \dots, t_n]$ be A -reduced and let $\lambda \in A_i$ for some i . If $f_\lambda = f(t_1, \dots, t_{i-1}, \lambda, t_{i+1}, \dots, t_n)$ vanishes everywhere on $A_1 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n$, then $t_i - \lambda$ divides f .*

Proof. Using Lemma 8.1.6 write $f = (t_i - \lambda)h + r$ with $\deg_{t_i} r < 1$, i.e., r can be considered as an element of $R[t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n]$. From Lemma 8.1.1 it follows that r is also A -reduced. Say $f_\lambda = r$ vanishes everywhere. Then by Lemma 8.1.5 we get $r = 0$, i.e., $t_i - \lambda$ divides f . \square

Theorem 8.2.6 (Punctured Combinatorial Nullstellensatz for Rings). *Let $f \in R[t_1, \dots, t_n]$ and let $A = A_1 \times \cdots \times A_n$ be a finite grid in R^n that satisfies Condition (D). Let B_1, \dots, B_n be nonempty subsets of A_1, \dots, A_n , respectively, and put $B = B_1 \times \cdots \times B_n$. For $i \in \{1, \dots, n\}$, define $\varphi_i = \prod_{\lambda \in A_i} (t_i - \lambda)$. If f vanishes on all points of $A \setminus B$ and does not vanish on at least one point of B , then there exist polynomials g_1, \dots, g_n such that $f = \sum g_i \varphi_i + r$ where $r \in \mathcal{R}_A$ is a non-zero multiple of the polynomial $\prod_{i=1}^n \prod_{\lambda \in A_i \setminus B_i} (t_i - \lambda)$ and $\deg g_i \leq \deg f - \deg \varphi_i$ for all i . Moreover, $\deg f \geq \deg r \geq \sum (|A_i| - |B_i|)$.*

Proof. Apply grid reduction on f with respect to A to get $f = \sum g_i \varphi_i + r$, such that $\deg g_i \leq \deg f - \deg \varphi_i$ for all i , and $\deg f \geq \deg r$. Then, we just have to show that $\prod_{i=1}^n \prod_{\lambda \in A_i \setminus B_i} (t_i - \lambda)$ divides r .

Let $\lambda \in A_1 \setminus B_1$. By the given condition on f , the polynomial $r(\lambda, t_2, \dots, t_n)$ vanishes everywhere on $A_2 \times \cdots \times A_n$. Now from Lemma 8.2.5 it follows that $t_1 - \lambda$ divides r . Write $r = (t_1 - \lambda)r'$. Say $\lambda' \in A_1 \setminus B_1$ with $\lambda' \neq \lambda$, then we see that $r(\lambda', t_2, \dots, t_n) = (\lambda' - \lambda)r'(\lambda', t_2, \dots, t_n)$ vanishes everywhere on $A_2 \times \cdots \times A_n$, and hence is equal to 0 by Lemma 8.2.5. Since $\lambda' - \lambda$ is not a zero divisor, the polynomial $r'(\lambda', t_2, \dots, t_n)$ must be equal to 0, which implies that $(t_1 - \lambda')$ divides r' . Continuing in this manner, we get that $\psi(t_1) = \prod_{\lambda \in A_1 \setminus B_1} (t_1 - \lambda)$ divides r .

Now we show that $\psi(t_2) = \prod_{\lambda \in A_2 \setminus B_2} (t_2 - \lambda)$, which also divides r by a similar reasoning, must in fact divide $r/\psi(t_1)$. Write $r = \psi(t_1)r_1$. Take a $\lambda \in A_2 \setminus B_2$. Then $r(t_1, \lambda, t_3, \dots, t_n) = \psi(t_1)r_1(t_1, \lambda, t_3, \dots, t_n)$ is equal to the zero polynomial by Lemmas 8.2.5 and 8.1.6. Since $\psi(t_1)$ is a monic non-zero polynomial, the polynomial $r_1(t_1, \lambda, t_3, \dots, t_n)$ must be equal to 0 by Lemma 8.1.2, and hence $t_2 - \lambda$ divides r_1 . Therefore, $\psi(t_1)\psi(t_2)$ divides r . Continuing in this manner, we get the result. \square

Note that Lemma 8.2.1, and hence all of its consequences, also follow from Theorem 8.2.6 by taking all B_i 's to be singleton.

²In fact, the same proof gives us Lason's generalization of Combinatorial Nullstellensatz [?]

³It's a variation in the sense that we do not take multiplicities into account and we work over grids satisfying Condition (D) over a ring.

8.3. Punctured Chevalley-Warning

From Theorem 8.2.6 over fields, we obtain a new generalization of the Chevalley-Warning theorem which we call *Punctured Chevalley-Warning Theorem*.

Theorem 8.3.1. *Let $A_1, \dots, A_n, B_1, \dots, B_n$ be nonempty finite subsets of \mathbb{F}_q such that for all i we have $B_i \subseteq A_i$. Put $A = A_1 \times \dots \times A_n$ and $B = B_1 \times \dots \times B_n$. Let $f_1, \dots, f_r \in \mathbb{F}_q[t_1, \dots, t_n]$ be such that $(q-1) \sum_{j=1}^r \deg f_j < \sum (|A_i| - |B_i|)$ and let $Z_A = \{x \in \prod A_i \mid \forall j, f_j(x) = 0\}$. If $Z_A \cap B \neq \emptyset$, then $Z_A \cap (A \setminus B) \neq \emptyset$.*

Proof. Let $f = \prod_{j=1}^r (1 - f_j^{q-1})$. Then we are given that $\deg f = (q-1) \sum \deg f_j < \sum (|A_i| - |B_i|)$. Say $Z_A \cap B \neq \emptyset$ and $Z_A \cap (A \setminus B) = \emptyset$. Then f vanishes everywhere on A , except at some point of B . Therefore, we get the contradiction from Theorem 8.2.6, which says that $\deg f = (q-1) \sum \deg f_j$ must be at least $\sum (|A_i| - |B_i|)$. \square

Again by taking B_i 's to be singletons we get Theorem 8.2.3 from Theorem 8.3.1. The following two direct corollaries of Theorem 8.3.1 might be of independent interest.

Corollary 8.3.2. *Let q be a prime power and let f_1, \dots, f_r be polynomials in $\mathbb{F}_q[t_1, \dots, t_n]$, with $\sum \deg f_i < n(q-2)/(q-1) = n - n/(q-1)$. If f_1, \dots, f_r have a common zero in $\{0, 1\}^n$, then they have a common zero in $\mathbb{F}_q^n \setminus \{0, 1\}^n$.*

Corollary 8.3.3. *Let q be a prime power and s a positive integer and let $\mathbb{F}_q \subset \mathbb{F}_{q^s}$ be a field extension. Let f_1, \dots, f_r be polynomials in $\mathbb{F}_q[t_1, \dots, t_n]$, with $\sum \deg f_i < n(q^s - q)/(q^s - 1) = n - n/(1 + q + \dots + q^{s-1})$. If f_1, \dots, f_r have a common zero in $\mathbb{F}_{q^s}^n$, then they have a common zero in $\mathbb{F}_{q^s}^n \setminus \mathbb{F}_q^n$.*

Remark. In the setting of Corollary 8.3.3, let N_s denote the number of common zeros of the polynomials in $\mathbb{F}_{q^s}^n$. Then the result gives a condition on the sum of degrees which ensures that $N_s > N_1$.

9. Alon-Füredi Theorem

9.1. Introduction

In [3], Alon and Füredi solved a problem posed by Bárány (based on a paper of Komjáth [99]) of finding the minimum number of hyperplanes required to cover all points of the hypercube $\{0, 1\}^n \subseteq F^n$ *except one*. One such covering is given by n hyperplanes defined by the zeros of the polynomials $t_1 - 1, t_2 - 1, \dots, t_n - 1$. Alon and Füredi proved that n is in fact the minimum number. They then generalized this result to all finite grids $A = \prod_{i=1}^n A_i \subseteq F^n$, showing that the minimum number of hyperplanes required to cover all points of A except one is $\sum_{i=1}^n (|A_i| - 1)$.¹

There is also a quantitative refinement: as we vary over families of d hyperplanes which do not cover all points of A , what is the minimum number of points that are missed? To answer this, Alon and Füredi proved the following result.

Theorem 9.1.1 (Alon-Füredi Theorem [3, Theorem 5]). *Let F be a field, let $A = \prod_{i=1}^n A_i \subseteq F^n$ be a finite grid, and let $f \in F[t]$ be a polynomial which does not vanish on all points of A . Then $f(x) \neq 0$ for at least $\min\{\prod y_i\}$ elements $x \in A$, where the minimum is taken over all positive integers $y_i \leq |A_i|$ with $\sum_{i=1}^n y_i = \sum_{i=1}^n |A_i| - \deg f$. More concisely (see Section 9.2 for the notation)*

$$|\mathcal{U}_A(f)| \geq \mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - \deg f).$$

Several proofs of Theorem 9.1.1 have been given. The original argument in [3] involves constructing auxiliary polynomial functions of low degree via linear algebra. A second proof was given by Ball and Serra as an application of Theorem 8.2.6 over fields. Recently, López, Rentería-Márquez and Villarreal gave a proof of Alon-Füredi in its coding theoretic formulation [107] (this will be discussed in Section 9.6). In [83], Geil showed that the minimum distance of generalized Reed-Muller codes can be determined easily using Gröbner basis theory [83, Theorem 2]. This technique was then used by Carvalho to give another proof of Theorem 9.1.1 when F is a finite field, [45, Proposition 2.3], which is in fact a special case of an earlier result by Geil and Thomsen [84, Proposition 5] (take all weights equal to 1).

In [50], Clark generalized the Alon-Füredi Theorem by replacing the field F by an arbitrary ring R , under the assumption that the finite grid A satisfies Condition (D), which as we have seen in Chapter 8 is exactly what is needed for polynomial functions on A to behave as they do in the case of a field. His proof adapts that of Ball and Serra.

¹This result follows directly from Lemma 8.2.1 by associating the hyperplanes to linear polynomials that define them and then taking the product of these polynomials.

Clark, Forrow and Schmitt [51] used Alon-Füredi to obtain a restricted variable generalization of a theorem of Warning [139] which gives a lower bound on the number of zeros of a system of polynomials over a finite field. Alon-Füredi gives a lower bound on *non-zeros*, but over a finite field \mathbb{F}_q , we have Chevalley's trick: $f(x) = 0 \iff 1 - f(x)^{q-1} \neq 0$ (see the proof of Theorem 8.2.2). This work also gave several combinatorial applications of this lower bound on restricted variable zero sets.

In this chapter, we will revisit the Alon-Füredi Theorem and give direct combinatorial applications (i.e., not of Chevalley-Warning type). We will also prove the following generalization of this result which, along with the total degree, takes the degrees of the polynomial in individual variables in account as well.

Theorem 9.1.2 (Generalized Alon-Füredi Theorem). *Let R be a ring and let A_1, \dots, A_n be nonempty finite subsets of R that satisfy Condition (D). For $i \in \{1, \dots, n\}$, let $a_i = |A_i|$ and let b_i be an integer such that $1 \leq b_i \leq a_i$. Let $f \in R[t]$ be a non-zero polynomial such that $\deg_{t_i} f \leq a_i - b_i$ for all $i \in \{1, \dots, n\}$. Let $\mathcal{U}_A = \{x \in A \mid f(x) \neq 0\}$ where $A = A_1 \times \dots \times A_n \subseteq R^n$. Then we have (see Section 9.2 for the notation)*

$$|\mathcal{U}_A| \geq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - \deg f).$$

Moreover, for any such R, A_1, \dots, A_n and integers b_1, \dots, b_n , we can construct a polynomial f which meets this bound.

Remark. Some justification is required for calling Theorem 9.1.2 a “Generalized” Alon-Füredi Theorem. If $f \in F[t_1, \dots, t_n]$ is a polynomial that does not vanish on all points of a finite grid A , then by Lemma 8.1.7, there exists a unique non-zero A -reduced polynomial $r_A(f)$ which takes the same values on A as f . We can use the polynomial $r_A(f)$, whose degree is at most $\deg f$, to show that Theorem 9.1.1 follows from Theorem 9.1.2 (we need Lemma 9.2.3 as well).

In Section 9.4 we relate the Generalized Alon-Füredi Theorem to work of DeMillo-Lipton, Schwartz and Zippel. We find in particular that Alon-Füredi implies the result which has become known as the “Schwartz-Zippel Lemma”. In fact, the original result of Zippel (and earlier, DeMillo-Lipton) is a bit different and not implied by Alon-Füredi (cf. Example 9.4). However, it is implied by Generalized Alon-Füredi, and this was one of our motivations for strengthening Alon-Füredi as we have. In Section 9.5 we discuss multiplicity enhancements in the sense of [78].

The Alon-Füredi Theorem has a natural coding theoretic interpretation (see Section 9.6) as it computes the minimum Hamming distance of the affine grid code $\text{AGC}_d(A)$, an F -linear code of length $|A|$. In this way Alon-Füredi turns out to be the restricted variable generalization of a much older result in the case $A_i = F = \mathbb{F}_q$, the Kasami-Lin-Peterson Theorem, which computes the minimum Hamming distance of generalized Reed-Muller codes. We will show that the Generalized Alon-Füredi Theorem is equivalent to the computation of the minimum Hamming distance of a more general class of R -linear codes. These generalized affine grid codes have larger distance (though also smaller dimension) than the standard ones, so they may turn out to be useful.

In Section 9.7, we pursue applications to finite geometry. We begin by revisiting and slightly sharpening the original result of Alon-Füredi on hyperplane coverings. This nat-

usually leads us to partial covers and blocking sets in affine and projective geometries over \mathbb{F}_q . Applying Alon-Füredi and projective duality we get a new upper bound, Theorem 9.7.6, on the number of hyperplanes which do not meet a k -element subset of $\text{AG}(n, q)$. From this result the classical theorems of Jamison-Brouwer-Schrijver on affine blocking sets and Blokhuis-Brouwer on essential points of projective blocking sets follow as corollaries. We are also able to strengthen a recent result of Dodunekov, Storme and Van de Voorde.

When combined with the work done by Clark, Forrow and Schmitt in [50, 51], our work demonstrates that the Alon-Füredi theorem, much like the Combinatorial Nullstellensatz, is a fundamental result on polynomials with connections to various important topics in mathematics. Thus we hope to convince the reader that this is an important tool to possess for those working in areas where polynomial methods might be successful.

9.2. Balls in Bins

Let $a_1, \dots, a_n \in \mathbb{Z}^+$. Consider n bins A_1, \dots, A_n such that A_i can hold up to a_i balls for every $i \in \{1, \dots, n\}$. For $N \in \mathbb{Z}^+$ with $n \leq N \leq \sum_{i=1}^n a_i$, we define a *distribution* of N balls in these n bins to be an n -tuple $y = (y_1, \dots, y_n) \in (\mathbb{Z}^+)^n$ with $y_i \leq a_i$ for all $i \in \{1, \dots, n\}$ and $\sum_{i=1}^n y_i = N$. For a distribution y of N balls in n bins, we put $P(y) = \prod_{i=1}^n y_i$. For $n \leq N \leq \sum_{i=1}^n a_i$ we define $\mathbf{m}(a_1, \dots, a_n; N)$ to be the minimum value of $P(y)$ as y ranges over all such distributions of N balls in n bins. For $N < n$ we define $\mathbf{m}(a_1, \dots, a_n; N) = 1$.

Without loss of generality we may assume $a_1 \geq \dots \geq a_n$. We define the *greedy distribution* $y_G = (y_1, \dots, y_n)$ as follows: first place one ball in each bin; then place the remaining balls into bins from left to right, filling each bin completely before moving on to the next bin, until we run out of balls. Then an easy argument shows the following.

Lemma 9.2.1. *Let $n \in \mathbb{Z}^+$, and let $a_1 \geq \dots \geq a_n$ be positive integers. Let $N \in \mathbb{Z}$ with $n \leq N \leq a_1 + \dots + a_n$.*

(1) *We have*

$$\mathbf{m}(a_1, \dots, a_n; N) = P(y_G) = y_1 \cdots y_n.$$

(2) *Suppose $a_1 = \dots = a_n = a \geq 2$. Then*

$$\mathbf{m}(a, \dots, a; N) = (r+1)a^{\lfloor \frac{N-n}{a-1} \rfloor},$$

where $r \equiv N - n \pmod{a-1}$ and $0 \leq r < a-1$.

(3) *For all non-negative integers k , we have*

$$\mathbf{m}(2, \dots, 2; 2n - k) = 2^{n-k}.$$

(4) *Let $n, a_1, \dots, a_n \in \mathbb{Z}^+$ with $a_1 \geq \dots \geq a_n$. Let $N \in \mathbb{Z}$ be such that $N - n = \sum_{i=1}^j (a_i - 1) + r$ for some $j \in \{0, \dots, n\}$ and some r satisfying $0 \leq r < a_{j+1}$. Then*

$$\mathbf{m}(a_1, \dots, a_n; N) = (r+1) \prod_{i=1}^j a_i.$$

Proof. Parts (1) through (3) are [51, Lemma 2.2]. (4) After placing one ball in each bin we are left with $N - n$ balls. The greedy distribution is achieved by filling the first j bins entirely and then putting r balls in bin $j + 1$. Note that (2) is a special case of (4). \square

In every distribution $y = (y_1, \dots, y_n)$ we need $y_i \geq 1$ for all $i \in \{1, \dots, n\}$; i.e., we must place at least one ball in each bin. So it is reasonable to think of the bins coming *prefilled* with one ball each, and then our task is to distribute the $N - n$ remaining balls so as to minimize $P(y)$. The concept of prefilled bins makes sense more generally: given any $b_1, \dots, b_n \in \mathbb{Z}$ with $1 \leq b_i \leq a_i$, we may consider the scenario in which the i -th bin comes prefilled with b_i balls. If $\sum_{i=1}^n b_i \leq N \leq \sum_{i=1}^n a_i$, we may restrict to distributions $y = (y_1, \dots, y_n)$ of N balls into bins of sizes a_1, \dots, a_n such that $b_i \leq y_i \leq a_i$ for all $i \in \{1, \dots, n\}$ and then we put

$$\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) = \min P(y),$$

where the minimum ranges over this restricted set of distributions. For $N < \sum_{i=1}^n b_i$ we define $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) := \prod_{i=1}^n b_i$.

Lemma 9.2.2. *We have $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) = \prod_{i=1}^n b_i \iff N \leq \sum_{i=1}^n b_i$.*

Proof. If $N \leq \sum_{i=1}^n b_i$ then $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) = \prod_{i=1}^n b_i$ by definition unless $N = \sum_{i=1}^n b_i$, and this case is immediate: we have exactly enough balls to perform the prefilling. If $N > \sum_{i=1}^n b_i$, then $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N)$ is the minimum over a set of integers each of which is strictly greater than $\prod_{i=1}^n b_i$. \square

In this prefilled context, the greedy distribution y_G is defined by starting with the bins prefilled with b_1, \dots, b_n balls and then distribute the remaining balls from left to right, filling each bin completely before moving on to the next bin. One sees, for example by adapting the argument of [51, Lemma 2.2], that

$$\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) = P(y_G)$$

when we also have $b_1 \geq \dots \geq b_n$. But this may not hold in general, as the following example shows.

Example. Let $n = 2$, $a_1 = 4$, $a_2 = 3$, $b_1 = 1$, $b_2 = 2$, $N = 4$. Then $P(y_G) = 4$ but $\mathbf{m}(4, 3; 1, 2; 4) = 3$ achieved by the distribution $(1, 3)$.

In general, we do not know a simple description of $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N)$. In practice, it can be computed using dynamic programming. The following properties of this combinatorial function will play an important role in our proof.

Lemma 9.2.3. *Let $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{Z}^+$ with $1 \leq b_i \leq a_i$ for all $i \in \{1, \dots, n\}$. Let $N_1, N_2 \in \mathbb{Z}$ such that $N_1 \leq N_2 \leq \sum_{i=1}^n a_i$. Then*

$$\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N_1) \leq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N_2).$$

Proof. If $N_2 \leq \sum b_i$, then both these quantities are equal by Lemma 9.2.2. For $N_1 \leq \sum b_i < N_2$, we have strict inequality. So assume that $\sum b_i < N_1 < N_2 \leq \sum a_i$. Let y be a distribution of N_2 balls for which $P(y) = \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N_2)$. From this distribution we can remove $N_2 - N_1 \geq 1$ balls without violating any conditions, which gives us a new distribution y' of N_1 balls. We must have $P(y') < P(y)$ and hence $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N_1) \leq P(y') < P(y) = \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N_2)$. \square

Lemma 9.2.4. Let $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{Z}^+$ with $1 \leq b_i \leq a_i$ for all $i \in \{1, \dots, n\}$. Let $k \in \mathbb{Z}$ such that $b_n \leq k \leq a_n$. If

$$b_1 + \dots + b_{n-1} \leq N - k \leq a_1 + \dots + a_{n-1}$$

for some $N \in \mathbb{Z}$, then

$$k \cdot \mathbf{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; N - k) \geq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N).$$

Proof. Let $y' = (y_1, \dots, y_{n-1})$ be a distribution of $N - k$ balls in the first $n - 1$ bins. Then $y = (y_1, \dots, y_{n-1}, k)$ is a distribution of N balls in n bins with the last bin getting k balls. Therefore,

$$\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) \leq P(y) = k \cdot P(y').$$

Since this holds for all such distributions y' , we get

$$\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N) \leq k \cdot \mathbf{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; N - k). \quad \square$$

9.3. Proof of Generalized Alon-Füredi Theorem

We go by induction on n .

Base Case: Let $f \in R[t_1]$ be a nonzero polynomial. Suppose f vanishes precisely at the distinct elements x_1, \dots, x_k of A_1 . By Corollary 8.1.4 we have

$$|\mathcal{U}_A(f)| = a_1 - k \geq a_1 - \deg f,$$

which is the conclusion of the Generalized Alon-Füredi Theorem in this case.

Induction Step: Suppose $n \geq 2$ and the result holds for $n - 1$. Write

$$f(t_1, \dots, t_n) = \sum_{i=0}^{d_n} f_i(t_1, \dots, t_{n-1}) t_n^i,$$

so that $d_n = \deg_{t_n} f$ is the largest index i such that $f_i \neq 0$. Moreover we have $\deg f_{d_n} \leq \deg f - d_n$ and for all $i \in \{1, \dots, n - 1\}$, $\deg_{t_i} f_{d_n} \leq \deg_{t_i} f \leq a_i - b_i$. Put $A' = \prod_{i=1}^{n-1} A_i$. By the induction hypothesis, we have

$$\begin{aligned} |\mathcal{U}_{A'}(f_{d_n})| &\geq \mathbf{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; \sum_{i=1}^{n-1} a_i - \deg f_{d_n}) \\ &\geq \mathbf{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; \sum_{i=1}^{n-1} a_i - \deg f + d_n), \text{ by Lemma 9.2.3.} \end{aligned}$$

Let $x' = (x_1, \dots, x_{n-1}) \in \mathcal{U}_{A'}(f_{d_n})$. Then $f(x', t_n) = \sum_{i=0}^{d_n} f_i(x') t_n^i \in R[t_n]$ has degree $d_n \geq 0$ since its leading term $f_{d_n}(x') t_n^{d_n}$ is non-zero. Since A_n satisfies Condition (D),

$f(x', t_n)$ vanishes at no more than d_n points of A_n by Corollary 8.1.4, so there are at least $a_n - d_n$ elements $x_n \in A_n$ such that $(x', x_n) \in \mathcal{U}_A(f)$. Thus

$$|\mathcal{U}_A(f)| \geq (a_n - d_n) \cdot \mathfrak{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; \sum_{i=1}^{n-1} a_i - \deg f + d_n).$$

Since

$$\deg f \leq \sum_{i=1}^n \deg_{t_i} f = \sum_{i=1}^{n-1} \deg_{t_i} f + d_n$$

and thus

$$\sum_{i=1}^{n-1} b_i \leq \sum_{i=1}^{n-1} (a_i - \deg_{t_i} f) \leq \sum_{i=1}^{n-1} a_i - \deg f + d_n \leq \sum_{i=1}^{n-1} a_i,$$

we may apply Lemma 9.2.4 with $N = \sum_{i=1}^n a_i - \deg f$ and $k = a_n - d_n$, getting

$$\begin{aligned} (a_n - d_n) \cdot \mathfrak{m}(a_1, \dots, a_{n-1}; b_1, \dots, b_{n-1}; \sum_{i=1}^{n-1} a_i - \deg f + d_n) &\geq \\ \mathfrak{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - \deg f). \end{aligned}$$

We deduce that

$$|\mathcal{U}_A(f)| \geq \mathfrak{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - \deg f).$$

9.3.1. Sharpness of the Generalized Alon-Füredi Bound

Let d be an integer such that $0 \leq d \leq \sum_{i=1}^n (a_i - b_i)$. For any distribution $y = (y_1, \dots, y_n)$ of $\sum_{i=1}^n a_i - d$ balls in n bins with $b_i \leq y_i \leq a_i$, for all i choose a subset $S_i \subseteq A_i$ of cardinality $a_i - y_i$, and put²

$$f(\underline{t}) = \prod_{i=1}^n \prod_{x_i \in S_i} (t_i - x_i).$$

Then

$$\deg f = \sum_{i=1}^n (a_i - y_i) = d,$$

$$\forall i \in \{1, \dots, n\}, \quad \deg_{t_i} f = a_i - y_i \leq a_i - b_i$$

and

$$\mathcal{U}_A(f) = P(y) = \prod_{i=1}^n y_i.$$

Thus, for all finite grids $A = \prod_{i=1}^n A_i$ satisfying Condition (D) and all permissible values of $\deg_{t_1} f, \dots, \deg_{t_n} f$ and $\deg f$, there are instances of equality in the Generalized Alon-

²An empty product is understood to take the value 1.

Füredi Bound. The case $b_1 = \dots = b_n = 1$ yields the (known) sharpness of the Alon-Füredi Bound.

9.4. Connections with the Schwartz-Zippel Lemma

The following result originated in computer science literature related to the Polynomial Identity Testing (PIT) problem where the aim is to design an “efficient” way of determining whether a given multivariate polynomial over a field F is identically equal to the zero polynomial (see [124, Chapter 4] for a survey and the precise problem statement). It is now known as the *Schwartz-Zippel Lemma* after the independent works of Schwartz [122] and Zippel [143].

Theorem 9.4.1 (Schwartz-Zippel Lemma). *Let R be an integral domain and let $S \subseteq R$ be finite and nonempty. Let $f \in R[t_1, \dots, t_n]$ be a nonzero polynomial. Then*

$$|Z_{S^n}(f)| \leq (\deg f)|S|^{n-1}.$$

This Lemma has a curious history, as recorded in the blog post of Lipton [106]. An earlier paper due to De Millo and Lipton [73] proves a similar result which is in fact sufficient for the applications to the PIT problem. This result of DeMillo and Lipton is also what Zippel proved in [143].

Theorem 9.4.2 (DeMillo-Lipton-Zippel Theorem). *Let R be an integral domain, let $f \in R[t_1, \dots, t_n]$, and let d be a positive integer such that $\deg_{t_i} f \leq d$ for all $i \in \{1, \dots, n\}$. Let S be a finite subset of R with more than d elements. Then*

$$|Z_{S^n}(f)| \leq |S|^n - (|S| - d)^n.$$

Schwartz proved Theorem 9.4.1 as a corollary of the following more general upper bound.

Theorem 9.4.3 (Schwartz Theorem [122, Lemma 1]). *Let R be an integral domain, $f = f_n \in R[t_1, \dots, t_n]$ be a nonzero polynomial and let $d_n = \deg_{t_n} f_n$. Let $f_{n-1} \in R[t_1, \dots, t_{n-1}]$ be the coefficient of $t_n^{d_n}$ in f_n . Let $d_{n-1} = \deg_{t_{n-1}} f_{n-1}$, and let $f_{n-2} \in R[t_1, \dots, t_{n-2}]$ be the coefficient of $t_{n-1}^{d_{n-1}}$ in f_{n-1} . Continuing in this manner we define for all $i \in \{1, \dots, n\}$ a polynomial $f_i \in R[t_1, \dots, t_i]$ with $\deg_{t_i} f_i = d_i$. Let $A = A_1 \times \dots \times A_n$ be a finite grid in R^n . Then*

$$|Z_A(f)| \leq |A| \sum_{i=1}^n \frac{d_i}{|A_i|}.$$

Interestingly, the DeMillo-Lipton-Zippel theorem does not imply the Schwartz-Zippel lemma, and neither is it implied by any of Schwartz’s results! We can see this via the following example.

Example. Let S be a finite subset of R containing 0, of size $s \geq 3$. Let $f = t_1 t_2 \in R[t_1, t_2]$. Then we have

$$|Z_{S^2}(f)| = 2s - 1.$$

DeMillo-Lipton-Zippel gives

$$|Z_{S^2}(f)| \leq s^2 - (s-1)^2 = 2s - 1.$$

Schwartz's Theorem and the Schwartz-Zippel lemma gives

$$|Z_{S^2}(f)| \leq s^2 \left(\frac{1}{s} + \frac{1}{s} \right) = 2s = (\deg f)|S|.$$

Thus neither Theorem 9.4.3, nor Theorem 9.4.1 imply Theorem 9.4.2. For the other direction, take $f = t_1 + t_2$. DeMillo-Lipton-Zippel gives $|Z_{S^2}(f)| \leq s^2 - (s-1)^2 = 2s - 1$, while the other results give $|Z_{S^2}(f)| \leq s$.

But, we can still relate these results using our main result, the Generalized Alon-Füredi Theorem, by showing that both Theorem 9.4.1 and Theorem 9.4.2 follow from Theorem 9.1.2. In fact, we will prove general versions of these results over arbitrary rings.

Theorem 9.4.4 (Generalized Schwartz-Zippel Lemma). *Let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite grid satisfying Condition (D), and suppose $|A_1| \geq \dots \geq |A_n|$. Let $f \in R[t_1, \dots, t_n]$ be a nonzero polynomial. Then*

$$|Z_A(f)| \leq (\deg f) \prod_{i=1}^{n-1} |A_i|.$$

Proof. If $\deg f \geq |A_n|$, then this holds trivially. So assume that $\deg f < |A_n|$. Then by Theorem 9.1.2, there are at least $\mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^{n-1} |A_i| + |A_n| - \deg f)$ points of A where f is non-zero. This expression is equal to $|A_1| \cdots |A_{n-1}|(|A_n| - \deg f)$ by Lemma 9.2.1, and hence f has at most $(\deg f) \prod_{i=1}^{n-1} |A_i|$ zeros in A . \square

Theorem 9.4.5 (Generalized DeMillo-Lipton-Zippel Theorem). *Let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite grid satisfying Condition (D), let $f \in R[t_1, \dots, t_n]$ be a nonzero polynomial and for $i \in \{1, \dots, n\}$ let $d_i = \deg_{t_i} f$. Assume that $0 \leq d_i < |A_i|$ for all $i \in \{1, \dots, n\}$. Then*

$$|Z_A(f)| \leq |A| - \prod_{i=1}^n (|A_i| - d_i).$$

Proof. For $i \in \{1, \dots, n\}$, put $a_i = |A_i|$ and $b_i = a_i - d_i$, so $1 \leq b_i \leq a_i$ for all i . Moreover $\deg f \leq \sum_{i=1}^n d_i$, so Theorem 9.1.2 gives

$$\begin{aligned} |\mathcal{U}_A| &\geq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - \deg f) \geq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n (a_i - d_i)) \\ &= \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n b_i) = \prod_{i=1}^n b_i = \prod_{i=1}^n (a_i - d_i). \end{aligned}$$

This proves the result as $|A| = |Z_A(f)| + |\mathcal{U}_A(f)|$. \square

Remark. (1) Using Corollary 8.1.4, one can easily adapt the original arguments of Theorems 9.4.1, 9.4.2 and 9.4.3 to prove their generalization over rings.

(2) A generalization of the Schwartz-Zippel Lemma over rings already appears in [6, Section 5] where the authors assume a stronger condition by taking subsets S of R

such that S is an integral domain in itself, and then considering the zeros of the polynomial in the grid S^n .

- (3) The case of the Schwartz-Zippel Lemma in which $R = S = \mathbb{F}_q$ is due to Ore [114]. Thus the Schwartz-Zippel Lemma may be viewed as a “Restricted Variable Ore Theorem”.

9.5. Multiplicity Enhancements

That one can assign to a zero of a polynomial a positive integer called *multiplicity* is a familiar concept in the univariate case. The definition of the multiplicity $m(f, x)$ of a multivariate polynomial $f \in R[t]$ at a point $x \in R^n$ may be less familiar, but the concept is no less useful. All of the main results considered thus far are upper bounds on $|Z_A(f)|$, the number of zeros of a polynomial f in a finite grid. By a *multiplicity enhancement* we mean the replacement of $|Z_A(f)|$ by $\sum_{x \in A} m(f, x)$ in such an upper bound. Here is the prototypical example: for a nonzero univariate polynomial f over a field F we have $\sum_{x \in F} m(f, x) \leq \deg f$.

Recently, multiplicity enhancements have become part of the polynomial method toolkit. In [78] Dvir, Kopparty, Saraf and Sudan gave a multiplicity enhancement of the Schwartz-Zippel Lemma. This was a true breakthrough with important applications in both combinatorics and theoretical computer science. In Section 9.4 we saw that the original work of Schwartz, DeMillo-Lipton and Zippel consists of more than the Schwartz-Zippel Lemma and gave some extensions of this work, in particular working over an arbitrary ring. So it is natural to consider multiplicity enhancements of these results. We do so here, giving a multiplicity enhancement of Theorem 9.4.3 over arbitrary rings R and thus also of Theorem 9.4.4. On the other hand the Alon-Füredi Theorem does not allow for a multiplicity enhancement (at least not in the precise sense described above), as we will see in Example 9.5.4.

In places our treatment closely follows that of [78]. We need to set things up over a ring, whereas they work over a field which makes things a bit more involved. Nevertheless, their work carries over verbatim much of the time, and when this is the case we state the result in the form we need it, cite the analogous result in [78] and omit the proof.

9.5.1. Hasse Derivatives

Let $R[t] = R[t_1, \dots, t_n]$. For $I = (i_1, \dots, i_n) \in \mathbb{N}^n$, put

$$\underline{t}^I = t_1^{i_1} \cdots t_n^{i_n}$$

and $|I| = \sum_{j=1}^n i_j = \deg \underline{t}^I$. Thus, $\{\underline{t}^I\}_{I \in \mathbb{N}^n}$ is an R -basis for $R[t]$. For $I = (i_1, \dots, i_n)$ and $J = (j_1, \dots, j_n)$ in \mathbb{N}^n , we put

$$\binom{I}{J} = \prod_{k=1}^n \binom{i_k}{j_k},$$

taking $\binom{i}{j} = 0$ if $j > i$. We say that $I \leq J$ if $i_k \leq j_k$ for all k . For $J \in \mathbb{N}^n$, let $D^J : R[\underline{t}] \rightarrow R[\underline{t}]$ be the unique R -linear map such that

$$D^J(\underline{t}^I) = \binom{I}{J} \underline{t}^{I-J}.$$

We have $D^J(\underline{t}^I) = 0$, unless $J \leq I$. Repeated application of the identity

$$t^n = (t - x + x)^n = \sum_{j=0}^n \binom{n}{j} x^{n-j} (t - x)^j$$

leads to the Taylor expansion: for $f \in R[\underline{t}]$ and $x \in R^n$,

$$f(\underline{t}) = \sum_{J \in \mathbb{N}^n} D^J(f)(x) (\underline{t} - x)^J. \quad (9.1)$$

Applying the automorphism $\underline{t} \mapsto \underline{t} + x$ gives the alternate form

$$f(\underline{t} + x) = \sum_{J \in \mathbb{N}^n} D^J(f)(x) \underline{t}^J.$$

These $D^J(f)$ were defined in [89] and are now called *Hasse derivatives*.

Proposition 9.5.1 ([78, Proposition 2.3]). *Let $f \in R[\underline{t}]$, and let $I, J \in \mathbb{N}^n$.*

- (1) *If f is homogeneous of degree d , then $D^I(f)$ is homogeneous of degree $d - |I|$.*
- (2) *We have*

$$D^J(D^I(f)) = \binom{I+J}{I} D^{I+J}(f).$$

Lemma 9.5.2 (Leibniz's Rule). *Let $f, g, h \in R[\underline{t}]$ such that $f = gh$. Then for $i \in \mathbb{N}$, we have*

$$D^i(f) = \sum_{j=0}^i D^j(g) D^{i-j}(h).$$

Proof. It suffices to show that $D^i(t^m t^n) = \sum_{j=0}^i D^j(t^m) D^{i-j}(t^n)$. On the left hand side we have $\binom{m+n}{i} t^{m+n-i}$, while on the right hand side we have $(\sum_{j=0}^i \binom{m}{j} \binom{n}{i-j}) t^{m+n-i}$, which are equal by the Vandermonde's identity for binomial coefficients. Alternately, look at the corresponding coefficients in $f(t+x)$, and $g(t+x)h(t+x)$. \square

9.5.2. Multiplicities

Let $f \in R[\underline{t}]$ be nonzero and $x \in R^n$. The *multiplicity of f at x* , denoted $m(f, x)$, is the natural number m such that $D^J(f)(x) = 0$ for all J with $|J| < m$ and $D^J(f)(x) \neq 0$ for some J with $|J| = m$. We put $m(0, x) = \infty$ for all $x \in R^n$.

Lemma 9.5.3 ([78, Lemma 2.4]). *For $f \in R[\underline{t}]$, $x \in R^n$ and $I \in \mathbb{N}^n$, we have*

$$m(D^I(f), x) \geq m(f, x) - |I|.$$

Given a vector $\underline{f} = (f_1, \dots, f_k) \in R[t]^k$, we put $m(\underline{f}, x) = \min_{1 \leq j \leq k} m(f_j, x)$.

Proposition 9.5.4 ([78, Proposition 2.5]). *Let $X_1, \dots, X_n, Y_1, \dots, Y_\ell$ be independent indeterminates. Let $\underline{f} = (f_1, \dots, f_k) \in R[X_1, \dots, X_n]^k$ and let $\underline{g} = (g_1, \dots, g_n) \in R[Y_1, \dots, Y_\ell]^n$. We define $\underline{f} \circ \underline{g} \in R[Y_1, \dots, Y_\ell]^k$ to be $\underline{f}(g_1, \dots, g_n)$.*

(1) *For any $a \in R^\ell$ we have*

$$m(\underline{f} \circ \underline{g}, a) \geq m(\underline{f}, \underline{g}(a))m(\underline{g} - \underline{g}(a), a).$$

(2) *In particular, since $m(\underline{g} - \underline{g}(a), a) \geq 1$, we have*

$$m(\underline{f} \circ \underline{g}, a) \geq m(\underline{f}, \underline{g}(a)).$$

Corollary 9.5.5 ([78, Corollary 2.6]). *Let $f \in R[t]$ and let $a, b \in R^n$. Then for all $c \in R$ we have*

$$m(f(a + tb), c) \geq m(f, a + cb).$$

The following Lemma will help us generalize the fact that $\sum_{x \in F} m(f, x) \leq \deg f$ for $f \in F[t]$ to polynomials with coefficients in an arbitrary commutative rings with identity.

Lemma 9.5.6. *Let $f, g, h \in R[t]$ be nonzero polynomials such that $f = gh$ and let $x \in R$. If $g(x)$ is not a zero divisor, then $m(f, x) = m(h, x)$.*

Proof. For any nonnegative integer i , from Lemma 9.5.2 we have

$$D^i(f) = D^0(g)D^i(h) + \dots + D^i(g)D^0(h).$$

Therefore $m(f, x) \geq m(h, x)$, as we have $D^i(f)(x) = 0$ for all $i < m(h, x)$. Let $m := m(f, x) \geq 1$ (for $m = 0$ the result is easily proved). We now show that $D^i(h)(x) = 0$ for all $i < m$, thus proving $m(h, x) \geq m(f, x)$. For $i = 0$, we have $0 = f(x) = g(x)h(x)$, and since $g(x)$ is not a zero divisor, we must have $h(x) = D^0(h)(x) = 0$. Now $D^1(f)(x) = D^0(g)(x)D^1(h)(x) + D^1(g)(x)D^0(h)(x) = g(x)D^1(h)(x)$. Again, since $g(x)$ is not a zero divisor, $D^1(f)(x) = 0$ if and only if $D^1(h)(x) = 0$. Continuing in this way, at the i 'th step we have $D^j(h)(x) = 0$ for all $j < i$ and thus $D^i(f)(x) = g(x)D^i(h)(x)$. Since $g(x)$ is not a zero divisor and $i < m$, we see that $D^i(h)(x) = 0$. \square

Lemma 9.5.7. *Let R be a ring, and let $f \in R[t]$ be a polynomial of degree $d \geq 1$. Let $A = \{x_1, \dots, x_n\} \subseteq R$ be a finite set satisfying Condition (D). Then*

$$\sum_{x \in A} m(f, x) \leq d.$$

Proof. We have $(t - x_i)^{m(f, x_i)} \mid f$ for all $i \in \{1, \dots, n\}$. In particular we may write

$$f(t) = (t - x_1)^{m(f, x_1)} g_1(t),$$

so $\deg f \geq m(f, x_1)$ and we are done if $n = 1$. So suppose $n \geq 2$. Since A satisfies Condition (D), the element $(x_n - x_1)^{m(f, x_1)} \dots (x_n - x_{n-1})^{m(f, x_{n-1})}$ of R is not a zero divisor. Therefore, if we have $f(t) = (\prod_{i=1}^{n-1} (t - x_i)^{m(f, x_i)}) g_{n-1}(t)$, then $m(f, x_n) = m(g_{n-1}, x)$ by Lemma 9.5.6 and hence $(t - x_n)^{m(f, x_n)} \mid g_{n-1}(t)$. Thus by induction we get $\prod_{i=1}^n (t - x_i)^{m(f, x_i)} \mid f(t)$, which implies that $\sum_{x \in A} m(f, x) \leq \deg f$. \square

From the proof of the multiplicity enhanced Schwartz-Zippel Lemma in [78], we extract the following lemma.

Lemma 9.5.8 (DKSS Lemma). *Let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite subset satisfying Condition (D). Let $f \in R[t]$, and write*

$$f = \sum_{j=0}^{d_n} f_j(t_1, \dots, t_{n-1})t_n^j$$

with $f_{d_n} \neq 0$. Put $A' = \prod_{i=1}^{n-1} A_i$. For all $x' = (x_1, \dots, x_{n-1}) \in A'$, we have

$$\sum_{x \in A_n} m(f, (x', x)) \leq |A_n| m(f_{d_n}, x') + d_n.$$

Proof. Fix an $x' \in A'$. Choose $I' \in \mathbb{N}^{n-1}$ such that $|I'| = m(f_{d_n}, x')$ and $D^{I'}(f_{d_n})(x') \neq 0$. Put $I = I' \times \{0\} \in \mathbb{N}^n$. Then

$$D^I(f) = \sum_{j=0}^{d_n} D^{I'}(f_j)t_n^j,$$

so $D^I(f) \neq 0$. By Lemma 9.5.3, we have

$$m(f, (x', x)) \leq |I| + m(D^I(f), (x', x)) = m(f_{d_n}, x') + m(D^I(f), (x', x)).$$

Apply Corollary 9.5.5 to $D^I(f)$ with $a = (x', 0)$, $b = (0, 1)$ and $c = x$: we get

$$m(D^I(f), (x', x)) \leq m(D^I(f)(x', t_n), x).$$

Summing over $x \in A_n$ gives

$$\sum_{x \in A_n} m(f, (x', x)) \leq |A_n| m(f_{d_n}, x') + \sum_{x \in A_n} m(D^I(f)(x', t_n), x).$$

Since $I = I' \times \{0\}$, $D^I(f)(x', t_n)$ has degree d_n and thus Lemma 9.5.7 gives

$$\sum_{x \in A_n} m(D^I(f)(x', t_n), x) \leq d_n.$$

The result follows. □

9.5.3. Multiplicity Enhanced Theorems

Theorem 9.5.9 (Multiplicity Enhanced Schwartz Theorem). *Let R be a ring, let $A = \prod_{i=1}^n A_i \subseteq R^n$ be finite, nonempty and satisfy Condition (D), and let $f = f_n \in F[t_1, \dots, t_n]$ be a nonzero polynomial. Let $d_n = \deg_{t_n} f$, and let $f_{n-1} \in R[t_1, \dots, t_{n-1}]$ be the coefficient of $t_n^{d_n}$ in f_n . Let $d_{n-1} = \deg_{t_{n-1}} f_{n-1}$, and let $f_{n-2} \in R[t_1, \dots, t_{n-2}]$ be the coefficient of $t_{n-2}^{d_{n-2}}$ in f_{n-1} . Continuing in this manner we define for all $1 \leq i \leq n$ a polynomial*

$f_i \in R[t_i, \dots, t_n]$ with $\deg_{t_i} f_i = d_i$. Then

$$\sum_{x \in A} m(f, x) \leq |A| \sum_{i=1}^n \frac{d_i}{|A_i|}.$$

Proof. We prove this by induction on n . The case $n = 1$ is Lemma 9.5.7. Suppose the result holds for polynomials in $n - 1$ variables. Let $A' = \prod_{i=1}^{n-1} A_i$. Applying Lemma 9.5.8 and then the induction hypothesis, we get

$$\begin{aligned} \sum_{x \in A} m(f, x) &= \sum_{x' \in A'} \sum_{x \in A_n} m(f, (x', x)) \leq |A_n| \sum_{x' \in A'} m(f_{n-1}, x') + |A'| d_n \\ &\leq |A_n| |A'| \sum_{i=1}^{n-1} \frac{d_i}{|A_i|} + |A| \frac{d_n}{|A_n|} = |A| \sum_{i=1}^n \frac{d_i}{|A_i|}. \end{aligned}$$

□

Theorem 9.5.10 (Multiplicity Enhanced Schwartz-Zippel Lemma). *Let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite grid satisfying Condition (D), and suppose $|A_1| \geq \dots \geq |A_n|$. Let $f \in R[t_1, \dots, t_n]$ be a nonzero polynomial. Then*

$$\sum_{x \in A} m(f, x) \leq (\deg f) \prod_{i=1}^{n-1} |A_i|.$$

Proof. Define d_i 's and f_i 's as in Theorem 9.5.9. Then the coefficient of $t_1^{d_1} \dots t_n^{d_n}$ is non-zero in f , and hence $\sum_{i=1}^n d_i \leq \deg f$. Thus we have

$$\sum_{x \in A} m(f, x) \leq |A| \sum_{i=1}^n \frac{d_i}{|A_i|} \leq |A_1| \dots |A_{n-1}| \sum_{i=1}^n d_i \leq (\deg f) \prod_{i=1}^{n-1} |A_i|.$$

□

Remark. (1) When R is a field, Theorem 9.5.9 was proved by Geil and Thomsen [84, Theorem 5]. They also build closely on [78].

(2) Unlike most of the other results presented here, Theorem 9.5.9 is *not* claimed to be sharp in all cases. In fact, it is not always sharp, and Geil and Thomsen give significant discussion of this point including an algorithm which sometimes leads to an improved bound [84, Theorem 6] and further numerical exploration.

(3) When R is a field and $A_1 = \dots = A_n$, then Theorem 9.5.10 appears as Lemma 2.7 in [78].

9.5.4. A Counterexample

It is natural to ask whether Alon-Füredi holds in multiplicity enhanced form, i.e., whether the bound

$$|Z_A(f)| \leq |A| - \mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - \deg f)$$

could be improved to

$$\sum_{x \in A} m(f, x) \leq |A| - \mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - \deg f),$$

in Theorem 9.1.1. The following example shows that such an improvement does not always hold.

Example. Let $n = 2$ and let $S = A_1 = A_2$ be a finite subset of s elements in R which contains 0, and put $A = S \times S$. Let $d_1, d_2 \in \mathbb{Z}^+$ be such that $d_1, d_2 < s \leq d_1 + d_2$. Then $f = t_1^{d_1} t_2^{d_2}$ is A -reduced, and we have $\mathbf{m}(s, s; 2s - d_1 - d_2) = 2s - d_1 - d_2 - 1$ (by the greedy distribution). But

$$\sum_{x \in A} m(f, x) = s(d_1 + d_2) > s^2 - 2s + d_1 + d_2 + 1 = s^2 - \mathbf{m}(s, s; 2s - d_1 - d_2),$$

since $(s-1)(d_1+d_2) > (s-1)^2$ (we have assumed $d_1+d_2 > s-1$) and thus the multiplicity enhanced version of Alon-Füredi would not hold in this case.

Remark. Since $s(d_1 + d_2) > s(d_1 + d_2) - d_1 d_2 = s^2 - (s - d_1)(s - d_2)$, we can use the polynomial $t_1^{d_1} t_2^{d_2}$, for any $d_1, d_2 < s$ to show that the multiplicity enhanced version of DeMillo-Lipton-Zippel Theorem (see Theorem 9.4.5) does not hold in all cases either.

9.6. Connections with Coding Theory

A q -ary linear code of length n is simply a subspace C of \mathbb{F}_q^n . The most important parameter of a code is the *Hamming distance* $d(C) = \min\{d(x, y) \mid x, y \in C, x \neq y\}$ where $d(x, y)$ for $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{F}_q^n$ is equal to $|\{i \in \{1, \dots, n\} \mid x_i \neq y_i\}|$. For linear codes $d(C)$ is simply equal to $\min\{w(x) \mid x \in C\}$ where $w(x) = d(x, 0)$ is the Hamming weight of x . These notions can be defined over arbitrary fields F and rings R as well. We refer to [104] for more background on Coding Theory. In this section we will see how the Alon-Füredi Theorem and its generalization are essentially statements that give the Hamming distance of certain linear codes.

Over the finite field \mathbb{F}_q , we can consider the whole space \mathbb{F}_q^n as a finite grid. Then the set of reduced polynomials with respect to this grid (see Chapter 8) are simply called reduced polynomials. We denote the set of reduced polynomials by $\mathcal{P}(n, q)$; it is an \mathbb{F}_q -vector space of dimension q^n . The evaluation map gives an \mathbb{F}_q -linear isomorphism

$$E : \mathcal{P}(n, q) \rightarrow \mathbb{F}_q^{\mathbb{F}_q^n}, f \mapsto (x \in \mathbb{F}_q^n \mapsto f(x)).$$

Fixing an ordering $\alpha_1, \dots, \alpha_{q^n}$ of \mathbb{F}_q^n , this isomorphism allows us to identify each $f \in \mathcal{P}(n, q)$ with its *value table* $(f(\alpha_1), \dots, f(\alpha_{q^n}))$. For $d \in \mathbb{N}$ we denote by $\mathcal{P}_d(n, q)$ the set of all reduced polynomials of degree at most d . The dimension of $\mathcal{P}_d(n, q)$ is equal to the number of integer solutions to $e_1 + \dots + e_n \leq d$, $e_1, \dots, e_n \geq 0$. The evaluation map restricted to $\mathcal{P}_d(n, q)$ also gives an injective \mathbb{F}_q -linear map.

Definition. The set of all value tables of all polynomials in $\mathcal{P}_d(n, q)$ is called the d -th order generalized Reed-Muller code of length q^n , denoted by $\text{GRM}_d(n, q)$.

For $q = 2$, these codes were introduced and studied by Muller [111] and Reed [118]. These were generalized to arbitrary q by Kasami, Lin and Peterson [97] who also gave an explicit formula for the minimum distance of the generalized Reed-Muller codes. A systematic study of these codes in terms of the polynomial formulation was conducted by Delsarte, Goethals and MacWilliams in [71], where they also classified all the minimum weight codewords. We will prove the result of Kasami, Lin and Peterson [97, Theorem 5] using the Alon-Füredi Theorem, thus giving shorter and more elementary proof than the original. While (binary) Reed-Muller codes are mentioned by Alon and Füredi in [3] under Corollary 1, the connection between [3, Theorem 5] and generalized Reed-Muller codes is not explored.

Theorem 9.6.1 (Kasami-Lin-Peterson). *The minimum weight of the d -th order Generalized Reed-Muller code $\text{GRM}_d(n, q)$ is equal to $(q - b)q^{n-a-1}$ where $d = a(q - 1) + b$ with $0 < b \leq q - 1$.*

Proof. The minimum weight of $\text{GRM}_d(n, q)$ is equal to the least number of nonzero values taken by a nonzero reduced polynomial of degree at most d , which by Alon-Füredi and its sharpness is equal to $\mathbf{m}(q, \dots, q; nq - d)$. Moreover we have

$$(nq - d) - n = n(q - 1) - a(q - 1) - b = (n - a - 1)(q - 1) + q - 1 - b,$$

and

$$0 \leq q - 1 - b < q - 1,$$

so by Lemma 9.2.1 we have

$$\mathbf{m}(q, \dots, q; nq - d) = (q - b)q^{n-a-1}. \quad \square$$

The Generalized Alon-Füredi Theorem can also be stated in terms of coding theory. Let $A = \prod_{i=1}^n A_i$ be a finite grid in R^n satisfying Condition (D), with $a_i = |A_i|$ for $i \in \{1, \dots, n\}$. Given positive integers $b_i \leq a_i$ for all $i \in \{1, \dots, n\}$, and a natural number $d \leq \sum_{i=1}^n (a_i - b_i)$, we define the *generalized affine grid code* $\text{GAGC}_d(A; b_1, \dots, b_n)$ as the set of value tables of all polynomials $f \in R[t]$ with $\deg_{t_i} f \leq a_i - b_i$ for all i and $\deg f \leq d$ evaluated on A . We put

$$\text{AGC}_d(A) = \text{GAGC}_d(A; 1, \dots, 1)$$

and speak of *affine grid codes*. Then from Theorem 9.1.2 (similar to Theorem 9.6.1) it follows that:

Theorem 9.6.2. *The minimum weight of $\text{GAGC}_d(A; b_1, \dots, b_n)$ is $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - d)$.*

Affine grid codes were studied in [107] under the name of Affine Cartesian Codes, where they proved the following:

Theorem 9.6.3 ([107, Thm. 3.8]). *Let F be a field and $A = \prod_{i=1}^n A_i \subseteq F^n$ a finite grid with $|A_1| \geq \dots \geq |A_n| \geq 1$. Then the minimum weight of $\text{AGC}_d(A)$ is*

$$\begin{cases} |A_1| \cdots |A_{k-1}|(|A_k| - \ell) & \text{if } d \leq \sum_{i=1}^n (|A_i| - 1) - 1, \\ 1 & \text{if } d \geq \sum_{i=1}^n (|A_i| - 1), \end{cases}$$

where $k, \ell \in \mathbb{Z}$ are such that $d = \sum_{i=k+1}^n (|A_i| - 1) + \ell$, $k \in \{1, \dots, n\}$ and $\ell \in \{1, \dots, |A_k| - 1\}$.

Proof. The minimum weight of $\text{AGC}_d(A)$ is $\mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - d)$. So the result follows from Lemma 9.2.1, as the greedy distribution of

$$\sum_{i=1}^n |A_i| - d = \sum_{i=1}^{k-1} (|A_i| - 1) + (|A_k| - 1 - \ell) + n$$

balls is $(|A_1|, \dots, |A_{k-1}|, |A_k| - \ell, 1, \dots, 1)$. \square

Remark. (1) The paper [107] makes no mention of Alon-Füredi. Their proof of Theorem 9.6.3 is self-contained and thus gives a proof of Alon-Füredi with the balls in bins constant replaced by its explicit value $P(y_G)$. On the other hand it is longer than the other proofs of Alon-Füredi appearing in the literature.

(2) When $b_1 \geq \dots \geq b_n$, the greedy algorithm computes $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N)$ and we could give a similarly explicit description of $\text{GAGC}_d(A_1; b_1, \dots, b_n)$.

9.7. Blocking Sets and Hyperplane Coverings

9.7.1. Partial Coverings of Grids by Hyperplanes

By a *hyperplane* in R^n we mean a polynomial $H = c_1 t_1 + \dots + c_n t_n + r \in R[t]$ for which at least one c_i is not a zero-divisor. (Referring to the polynomial itself rather than its zero locus in R^n will make the discussion cleaner.) A family $\mathcal{H} = \{H_i\}_{i=1}^d$ covers $x \in R^n$ if $H_i(x) = 0$ for some $i \in \{1, \dots, n\}$; \mathcal{H} covers a subset $S \subseteq R^n$ if it covers every point of S , and \mathcal{H} partially covers S otherwise. For a family $\mathcal{H} = \{H_i\}_{i=1}^d$ of hyperplane in R^n , put

$$f_{\mathcal{H}} = \prod_{i=1}^d H_i.$$

Thus $f_{\mathcal{H}} \in R[t_1, \dots, t_n]$ is a polynomial of degree d . Clearly, if \mathcal{H} covers a point x , then $f_{\mathcal{H}}(x) = 0$. When R is an integral domain, the converse also holds, and then we can identify the points of R^n covered by \mathcal{H} with the zeros of the polynomial $f_{\mathcal{H}}$ that has degree $|\mathcal{H}|$. We now revisit the original combinatorial problem studied by Alon and Füredi, which is part (3) of the following theorem. However, our proof is via Theorem 9.1.1 instead of the approach used in [3].

Theorem 9.7.1. *Let R be an integral domain, let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite grid, and let $\mathcal{H} = \{H_i\}_{i=1}^d$ be family of hyperplanes in R^n .*

- (1) *If \mathcal{H} partially covers A , then \mathcal{H} fails to cover at least $\mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - d)$ points of A .*
- (2) *For all $d \in \mathbb{Z}^+$, there is a family of hyperplane $\{H_1, \dots, H_d\}$ where each H_i is of the form $t_{j_i} - x_i$ for some $j_i \in \{1, \dots, n\}$ and $x_i \in A_{j_i}$ such that this family covers all but exactly $\mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - d)$ points of A .*
- (3) *If \mathcal{H} covers all but exactly one point of A , then $d \geq \sum_{i=1}^n (|A_i| - 1)$.*

Proof. (1) As above, \mathcal{H} covers $x \in R^n$ iff $f_{\mathcal{H}}(x) = 0$. Now apply the Alon-Füredi Theorem to $f_{\mathcal{H}}$.

(2) The sharpness construction of Section 9.3.1 is precisely of this form.

(3) Say \mathcal{H} covers all points of A except one. Then

$$1 = \mathcal{U}_A(f_{\mathcal{H}}) \geq \mathbf{m}(|A_1|, \dots, |A_n|; \sum_{i=1}^n |A_i| - d).$$

Now if $d < \sum_{i=1}^n (|A_i| - 1)$, then $\sum_{i=1}^n |A_i| - d \geq n + 1$, and since $\mathbf{m}(|A_1|, \dots, |A_n|; n + 1) = 2$, we get a contradiction by using Lemma 9.2.3.

□

Remark. (1) The proof of Theorem 9.7.1(3) can clearly be adapted to any polynomial of degree d which vanishes on all points of the grid except one. In this manner we get an alternate proof of Lemma 8.2.1 using the Alon-Füredi theorem.

(2) Let $R = \mathbb{Z}/4\mathbb{Z}$. Then we have a family of 5 lines in R^2 which cover all points of R^2 except the origin, given by $\mathcal{H} = \{t_1 - 1, t_1 - 2, t_1 - 3, t_2 - 2, t_1 + 2t_2 - 2\}$. Therefore, Theorem 9.7.1 need not be true when R is not an integral domain. In fact, this example can be generalized to show that Theorem 9.7.1 does not hold for any finite grid A in R^n which does not satisfy Condition (D). When A does satisfy Condition (D), and R is not an integral domain, we do not know if the result is true.

We complement Theorem 9.7.1 by computing the minimum cardinality of a hyperplane covering of a finite grid (not necessarily satisfying Condition (D)) over a ring R .

Theorem 9.7.2. *Let $A = \prod_{i=1}^n A_i \subseteq R^n$ be a finite grid, and let $\mathcal{H} = \{H_i\}_{i=1}^d$ be a hyperplane covering of A . Then $d \geq \min\{|A_i| \mid i \in \{1, \dots, n\}\}$.*

Proof. First we observe that if A satisfies Condition (D) then the result is almost immediate: going by contraposition, if $d \leq |A_i| - 1$ for all $i \in \{1, \dots, n\}$ then $f_{\mathcal{H}}$ is nonzero and A -reduced, so it cannot vanish identically on A by Lemma 8.1.5. Now we give a non-polynomial method proof in the general case. Without loss of generality assume $|A_1| \geq \dots \geq |A_n|$. We claim that any hyperplane $H = \sum_{i=1}^n c_i t_i + g$ covers at most $\prod_{i=1}^{n-1} |A_i|$ points of A : this suffices, for then $d \geq |A_n|$.

Proof of claim: Fix $i \in \{1, \dots, n\}$ such that c_i is not a zero-divisor in R . Let $\pi : R^n \rightarrow R^{n-1}$ be the projection $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Then

$$A = \coprod_{x' = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \pi(A)} \{x_1\} \times \dots \times \{x_{i-1}\} \times A_i \times \{x_{i+1}\} \times \dots \times \{x_n\}$$

is a partition of A into $|\pi(A)| = \prod_{j \neq i} |A_j|$ nonempty subsets, each one of which meets H in at most one point since c_i is not a zero divisor. So

$$|(Z(H) \cap A)| \leq \prod_{j \neq i} |A_j| \leq \prod_{i=1}^{n-1} |A_i|.$$

□

Conjecture 9.7.3. *Let R be a ring, and let $A_1, \dots, A_n \subseteq R$ be nonempty (but possibly infinite). Let $\mathcal{H} = \{H_j\}_{j \in J}$ be a covering of the grid $A = \prod_{i=1}^n A_i$ by hyperplanes. Then $|J| \geq \min\{|A_i| \mid i \in J\}$.*

Remark. (1) For $i \in \{1, \dots, n\}$, let $B_i \subseteq A_i \subseteq R$. Then we need at least as many hyperplanes to cover $\prod_{i=1}^n A_i$ as we do to cover $\prod_{i=1}^n B_i$. Together with Theorem 9.7.2 it follows that in the setting of Conjecture 9.7.3 we need at least $\min(|A_i|, \aleph_0)$ hyperplanes, since if A is infinite then for every finite subgrid $B = B_1 \times \dots \times B_n$ of A we need at least $\min\{|B_i|\}$ hyperplanes to cover it. Thus Conjecture 9.7.3 holds when R is countable.

(2) When R is a field and $A = R^n$, Conjecture 9.7.3 is a case of [48, Theorem 3].

9.7.2. Partial Covers and Blocking Sets in Finite Geometries

The same ideas can be used to prove old and new results about projective and affine spaces over finite fields.

Let $\text{PG}(n, q)$ denote the n -dimensional projective space over \mathbb{F}_q and let $\text{AG}(n, q)$ denote the n -dimensional affine space over \mathbb{F}_q . The set $\text{AG}(n, q)$ comes equipped with a sharply transitive action of the additive group of \mathbb{F}_q^n and thus a choice of a point $x \in \text{AG}(n, q)$ induces an isomorphism $\text{AG}(n, q) \cong \mathbb{F}_q^n$. We will make such identifications without further comment.

A *partial cover* of $\text{PG}(n, q)$ is a set of hyperplanes that do not cover all the points. The points missed by a partial cover are called *holes*.

Theorem 9.7.4. *Let \mathcal{H} be a partial cover of $\text{PG}(n, q)$ of size $k \in \mathbb{Z}^+$. Then \mathcal{H} has at least $\mathbf{m}(q, \dots, q; nq - k + 1)$ holes.*

Proof. Let $H \in \mathcal{H}$. Then $\text{PG}(n, q) \setminus H \cong \text{AG}(n, q)$ so $\mathcal{H} \setminus H$ is a partial cover of \mathbb{F}_q^n by $k - 1$ hyperplanes. As in Theorem 9.7.1, there are at least $\mathbf{m}(q, \dots, q; nq - (k - 1))$ points not covered by \mathcal{H} . \square

Corollary 9.7.5. *If $0 \leq a < q$, a partial cover of $\text{PG}(n, q)$ of size $q + a$ has at least $q^{n-1} - aq^{n-2}$ holes.*

Proof. By Theorem 9.7.4 there are at least $\mathbf{m}(q, \dots, q; (n - 1)q - a + 1)$ holes. Since $0 \leq a < q$, the greedy distribution is $(q, \dots, q, q - a, 1)$, and the result follows. \square

Dodunekov, Storme and Van de Voorde have shown that a partial cover of $\text{PG}(n, q)$ of size $q + a$ has at least $q^{n-1} - aq^{n-2}$ holes if $0 \leq a < \frac{q-2}{3}$ [75, Theorem 17]. Corollary 9.7.5 gives an improvement in that the restriction on a is relaxed. They also show that if $a < \frac{q-2}{3}$ and the number of holes are at most q^{n-1} , then they are all contained in a single hyperplane. We cannot make any such conclusions from our arguments.

Projective duality yields a dual form of Theorem 9.7.4: k points in $\text{PG}(n, q)$ which do not meet all hyperplanes must miss at least $\mathbf{m}(q, \dots, q; nq - k + 1)$ of them. Thus:

Theorem 9.7.6. *Let S be a set of k points in $\text{AG}(n, q)$. Then there are at least $\mathbf{m}(q, \dots, q; nq - k + 1) - 1$ hyperplanes of $\text{AG}(n, q)$ which do not meet S .*

Proof. Add a hyperplane at infinity to get to the setting of $\text{PG}(n, q)$ and then apply the dual form of Theorem 9.7.4. \square

The general problem of the number of linear subspaces missed by a given set of points in $\text{PG}(n, q)$ is studied by Metsch in [109]. We wish to note that Theorem 9.7.6 gives the same bounds as Part a) of [109, Theorem 1.2] for the specific case when the linear subspaces are hyperplanes.

A *blocking set* in $\text{AG}(n, q)$ or $\text{PG}(n, q)$ is a set of points that meets every hyperplane. The union of the coordinate axes in \mathbb{F}_q^n yields a blocking set in $\text{AG}(n, q)$ of size $n(q - 1) + 1$. Doyen conjectured in a 1976 Oberwolfach lecture that $n(q - 1) + 1$ is the least possible size of a blocking set in $\text{AG}(n, q)$. A year later Jamison [93] proved this conjecture, and then an alternate (and simpler) proof by Brouwer and Schrijver [37] was found. We are in a position to give another proof.

Corollary 9.7.7 (Jamison-Brouwer-Schrijver). *The minimum size of a blocking set in $\text{AG}(n, q)$ is $n(q - 1) + 1$.*

Proof. Let $B \subseteq \text{AG}(n, q)$ be a blocking set of cardinality at most $n(q - 1)$. By Theorem 9.7.6, Lemma 9.2.1 and Lemma 9.2.3, there are at least

$$\mathbf{m}(q, \dots, q; nq - n(q - 1) + 1) - 1 = \mathbf{m}(q, \dots, q; n + 1) - 1 = 1$$

hyperplanes which do not meet B . \square

Turning now to $\text{PG}(n, q)$, every line is a blocking set. But classifying blocking sets that do not contain any line is one of the major open problems in finite geometry. For a survey on blocking sets in finite projective spaces, see [55, Chapter 3].

If $B \subseteq \text{PG}(n, q)$, $x \in B$ and H is a hyperplane in $\text{PG}(n, q)$, then H is a *tangent* to B through x if $H \cap B = \{x\}$. An *essential point* of a blocking set B in $\text{PG}(n, q)$ is a point x such that $B \setminus \{x\}$ is not a blocking set. A point x of B is essential if and only if there is a tangent hyperplane to B through x .

Theorem 9.7.8. *Let B be a blocking set in $\text{PG}(n, q)$ and x an essential point of B . There are at least $\mathbf{m}(q, \dots, q; nq - |B| + 2)$ tangent hyperplanes to B through x .*

Proof. Let H be a tangent hyperplane to B through x . Then $B' = B \setminus \{x\} \subseteq \text{PG}(n, q) \setminus H \cong \text{AG}(n, q)$. By Theorem 9.7.6 there are at least $\mathbf{m}(q, \dots, q; nq - |B| + 2) - 1$ hyperplanes in $\text{AG}(n, q)$ that do not meet B' . Since B is a blocking set all of these hyperplanes, when seen in $\text{PG}(n, q)$, must meet x . Thus there are at least $\mathbf{m}(q, \dots, q; nq - |B| + 2)$ tangent hyperplanes to B through x . \square

Corollary 9.7.9 (Blokhuis-Brouwer [26]). *Let B be a blocking set in $\text{PG}(2, q)$ of size $2q - s$. There are at least $s + 1$ tangent lines through each essential point of B .*

Proof. By Theorem 9.7.8, each essential point of B has at least

$$\mathbf{m}(q, q; 2q - (2q - s) + 2) = \mathbf{m}(q, q; s + 2)$$

tangent lines. Since $|B| = 2q - s < q^2 + q + 1 = |\text{PG}(2, q)|$, there exists $x \in \text{PG}(2, q) \setminus B$. There are $q + 1$ lines through x , so $2q - s = |B| \geq q + 1$. Thus $s + 1 \leq q$, so the greedy distribution is $(s + 1, 1)$ and $\mathbf{m}(q, q; s + 2) = s + 1$. \square

Corollary 9.7.10 ([75, Theorem 7]). *If $0 \leq a < q$, there are at least $q^{n-1} - aq^{n-2}$ tangent hyperplanes through each essential point of a blocking set of size $q + a + 1$ in $\text{PG}(n, q)$.*

Proof. By Theorem 9.7.8 and the proof of Corollary 9.7.5, each essential point of B has at least $\mathfrak{m}(q, \dots, q; nq - (q + a + 1) + 2) = q^{n-1} - aq^{n-2}$ tangent hyperplanes. \square

A. Nederlandstalige samenvatting

Deze thesis bestaat uit twee delen. In het eerste deel van de thesis werd de theorie van valuaties van schier veelhoeken, in combinatie met computerberekeningen (in GAP en SageMath), gebruikt om enerzijds nieuwe schier achthoeken te construeren en te bestuderen en anderzijds om karakterisatieresultaten te bekomen van schier veelhoeken in termen van bepaalde deelmeetkunden. In het tweede deel van de thesis wordt er dieper ingegaan op de polynomiale methode die toelaat om via allerhande manipulaties van polynomen bepaalde interessante resultaten uit de getallenleer, combinatoriek, eindige meetkunde, ... te bewijzen.

A.1. Deel 1 (Hoofdstukken 1 t.e.m. 6)

Een *schier veelhoek* is een punt-rechte meetkunde \mathcal{S} die aan de volgende eigenschappen voldoet:

- Elke twee verschillende punten zijn bevat in tenhoogste één rechte.
- De collineariteitsgraad Γ van \mathcal{S} heeft eindige diameter.
- Voor elk punt x en elke rechte L bestaat er een uniek punt op L dat het dichtst bij x gelegen is met betrekking tot de afstand in Γ .

Ingeval d de diameter van Γ is, wordt de schier veelhoek ook een *schier $2d$ -hoek* genoemd. Een schier veelhoek heeft *orde* (s, t) als elke rechte precies $s + 1$ punten bevat, en als elk punt bevat is in precies $t + 1$ rechten. Een schier $2d$ -hoek met $d \geq 2$ wordt een *veralgemeende $2d$ -hoek* genoemd als de volgende additionele eigenschappen voldaan zijn:

- Elk punt is incident met tenminste twee rechten.
- Als x en y twee punten zijn op afstand $i \in \{1, 2, \dots, d - 1\}$, dan heeft y een unieke buur die op afstand $i - 1$ van x gelegen is.

In het eerste deel van de thesis wordt veelvuldig gebruik gemaakt van volgende schier veelhoeken.

- $H(q, 1)$ stelt de veralgemeende zeshoek van orde $(q, 1)$ voor die afkomstig is van het Desarguesiaans projectief vlak $PG(2, q)$,
- $H(q)$ stelt de split Cayley veralgemeende zeshoek van orde (q, q) voor.
- $H(q)^D$ stelt de punt-rechte dual van $H(q)$ voor.
- $T(q, q^2)$ stelt de duale twisted triality hexagon van orde (q, q^2) voor.
- HJ stelt de Hall-Janko schier achthoek voor die geconstrueerd werd door Cohen [52].

Constructies van bovenvermelde veralgemeende zeshoeken kunnen gevonden worden in het boek [138].

Eén van de hoofdresultaten uit de thesis is de constructie in Hoofdstuk 5 van twee nieuwe schier achthoeken die een grote automorfismegroep hebben.

Theorem A.1.1. (a) Noem O_1 de punt-rechte meetkunde waarvan de punten de 4095 centrale involuties van de groep $G_1 := G_2(4):2$ zijn en waarvan de rechten alle triples $\{x, y, xy\}$ zijn, waarbij x en y twee commuterende centrale involuties van $G_2(4):2$ zijn waarvoor $[G_1 : N_{G_1}(\langle x, y \rangle)] \in \{1365, 13650\}$. Dan is O_1 een schier achthoek van orde $(2, 10)$.

(b) Noem S_1 de verzameling van alle rechten $\{x, y, z\}$ waarbij x en y twee commuterende involuties van $G_1 = G_2(4):2$ zijn waarvoor $[G_1 : N_{G_1}(\langle x, y \rangle)] = 1365$. Dan is S_1 een rechtenspread van O_1 . Als Q_1 de verzameling van alle quads van O_1 is, dan is de punt-rechte meetkunde \mathcal{H}_1 met puntenverzameling S_1 en rechtenverzameling Q_1 een veralgemeende zeshoek die isomorf is met de duale split Cayley hexagon $H(4)^D$, als we voor incidentie “het bevat zijn in” nemen.

(c) De automorfismegroep van O_1 is isomorf met $G_2(4):2$. The automorfismen van O_1 zijn precies de toevoegingen met elementen van $G_2(4):2$.

Theorem A.1.2. (a) Noem O_2 de punt-rechte meetkunde waarvan de punten de 315 centrale involuties van de groep $G_2 := L_3(4):2^2$ zijn en waarvan de rechten alle triples $\{x, y, xy\}$ zijn, waarbij x en y twee commuterende centrale involuties van $L_3(4):2^2$ zijn waarvoor $[G_2 : N_{G_2}(\langle x, y \rangle)] \in \{105, 420\}$. Dan is O_2 een schier achthoek van orde $(2, 4)$.

(b) Noem S_2 de verzameling van alle rechten $\{x, y, z\}$ waarbij x en y twee commuterende involuties van $G_2 = L_3(4):2^2$ zijn waarvoor $[G_2 : N_{G_2}(\langle x, y \rangle)] = 105$. Dan is S_2 een rechtenspread van O_2 . Als Q_2 de verzameling van alle quads van O_2 is, dan is de punt-rechte meetkunde \mathcal{H}_2 met puntenverzameling S_2 en rechtenverzameling Q_2 een veralgemeende zeshoek van orde $(4, 1)$, als we voor incidentie “het bevat zijn in” nemen.

(c) De automorfismegroep van O_2 is isomorf met $L_3(4):2^2$. The automorfismen van O_2 zijn precies de toevoegingen met elementen van $L_3(4):2^2$.

De schier achthoeken O_1 en O_2 worden de $G_2(4)$ en $L_3(4)$ schier achthoeken genoemd. In de thesis werd eveneens aangetoond dat O_2 ingebed kan worden in O_1 als een volle deelmeetkunde. Naast O_2 heeft de $G_2(4)$ schier achthoek O_1 nog andere interessante deelmeetkunden.

Theorem A.1.3. Onderstel dat H een maximale deelgroep is van $G_2(4):2$ isomorf met $J_2:2$. Stel dat X gelijk is aan de verzameling van de centrale involuties van $G_2(4):2$ die bevat zijn in H . Dan is X een deelruimte van O_1 , en de deelmeetkunde geïnduceerd op X door de rechten van O_1 die al hun punten in X hebben, is isomorf met met Hall-Janko schier achthoek HJ . Omgekeerd, wordt elke deelmeetkunde van O_1 isomorf met HJ op deze manier bekomen.

Met de ontdekking van de $G_2(4)$ schier achthoek O_1 , kunnen we nu volgende keten van schier veelhoeken neerschrijven:

$$H(2, 1) \subset H(2)^D \subset HJ \subset O_2.$$

Deze keten heeft sterke gelijkenissen met de Suzuki keten $L_3(2) \leq U_3(3) \leq J_2 \leq G_2(4) \leq Suz$ van enkelvoudige groepen. De volgende karakterisaties van de “Suzuki keten schier veelhoeken” werden bekomen in Hoofdstuk 6 van de thesis. De bewijzen steunen op de theorie van de valuaties van schier veelhoeken die ontwikkeld werd in [59, 64].

- Theorem A.1.4.** (a) *Op isomorfie na is de duale split Cayley veralgemeende zeshoek $H^D(2)$ de unieke schier zeshoek van orde $(2, 2)$ die $H(2, 1)$ bevat als een volle isometrisch ingebedde meetkunde.*
- (b) *Op isomorfie na is de Hall-Janko schier achthoek HJ de unieke schier achthoek van orde $(2, 4)$ die $H(2)^D$ bevat als een volle isometrisch ingebedde meetkunde.*
- (c) *Op isomorfie na is de $G_2(4)$ schier achthoek O_1 de unieke schier achthoek van orde $(2, 10)$ die HJ bevat als een volle isometrisch ingebedde meetkunde.*

Eén van de belangrijkste open problemen in de theorie van de veralgemeende veelhoeken is het al of niet bestaan van *half-oneindige veralgemeende veelhoeken*. Dit zijn veralgemeende veelhoeken van orde (s, t) waarbij $s \geq 2$ eindig en t oneindig is. Het volgende werd aangetoond in Hoofdstuk 4.

Theorem A.1.5. *Als $q \in \{2, 3, 4\}$, dan zijn er geen half-oneindige veralgemeende zeshoeken van orde (q, t) die $H(q)$ of $H(q)^D$ als volle deelmeetkunde bevatten.*

Het bewijs in het geval van een deelmeetkunde isomorf met $H(4)^D$ steunde op volgend resultaat uit Hoofdstuk 3, dat werd bekomen met behulp van computerberekeningen.

Theorem A.1.6. *De duale split Cayley hexagon $H(4)^D$ heeft geen afstands-2-ovoiden, d.w.z. geen puntenverzamelingen die elke rechte snijden in een singleton.*

Hoofdstuk 3 is in het algemeen gewijd aan het computationeel aspect van de thesis, met beschrijving van algoritmen en computationale technieken. De volgende resultaten over veralgemeende zeshoeken werden eveneens in Hoofdstuk 4 aangetoond.

Theorem A.1.7. *Als $q \in \{2, 4\}$, dan zijn er geen veralgemeende zeshoeken die $H(q)$ als een eigenlijke volle deelmeetkunde bevatten.*

Theorem A.1.8. *Elke schier zeshoek met drie punten per rechte die een isometrisch ingebedde volle deelmeetkunde heeft isomorf met $H(2)^D$, is isomorf met ofwel $H(2)^D$ ofwel $T(2, 8)$.*

A.2. Deel 2 (Hoofdstukken 7 t.e.m. 9)

Als R een ring is, dan wordt elke verzameling van de vorm $A = A_1 \times \cdots \times A_n$, waarbij A_i een niet-ledige eindige deelverzameling van R is, een *eindig rooster* van R^n genoemd. We zeggen dat zo'n rooster aan *Conditie (D)* voldoet als voor elke $i \in \{1, 2, \dots, n\}$ en alle $\alpha, \beta \in A_i$ met $\alpha \neq \beta$, het element $\alpha - \beta$ geen nuldeler is van R . Er kan aangetoond worden dat als A aan Conditie (D) voldoet, dan bestaat er voor elke veelterm $f \in R[t_1, \dots, t_n]$ een unieke veelterm $r_A(f) \in R[t_1, \dots, t_n]$ zodat $\deg_{t_i} r_A(f) \leq |A_i| - 1$ voor alle $i \in \{1, 2, \dots, n\}$ en $f(x) = r_A(f)(x)$ voor alle $x \in A$. Hiervan gebruik makend, werden in Hoofdstuk 8 van de thesis bewijzen gegeven van de “Combinatorial Nullstellensatz” en de “Punctured Combinatorial Nullstellensatz”. In Hoofdstuk 8 van de thesis werd eveneens volgende nieuwe veralgemening van het klassieke Chevalley-Warning theorema bewezen.

Theorem A.2.1. *Onderstel dat $A_1, \dots, A_n, B_1, \dots, B_n$ niet-ledige deelverzamelingen zijn van \mathbb{F}_q zodat $B_i \subseteq A_i$ voor alle $i \in \{1, 2, \dots, n\}$. Stel $A = A_1 \times \cdots \times A_n$ en $B = B_1 \times \cdots \times B_n$. Onderstel eveneens dat $f_1, \dots, f_r \in \mathbb{F}_q[t_1, \dots, t_n]$ zodat $(q-1) \sum_{j=1}^r \deg f_j < \sum (|A_i| - |B_i|)$ en definieer $Z_A = \{x \in \prod A_i \mid \forall j f_j(x) = 0\}$. Dan geldt dat $Z_A \cap (A \setminus B) \neq \emptyset$ als $Z_A \cap B \neq \emptyset$.*

In Hoofdstuk 9 van de thesis, werd volgende veralgemening van het Alon-Füredi theorema bewezen, en vervolgens toegepast in verschillende gebieden zoals de codeertheorie, theoretische computerwetenschappen en eindige meetkunde.

Theorem A.2.2 (Veralgemeend Alon-Füredi Theorema). *Onderstel dat $A = A_1 \times \cdots \times A_n \subseteq R^n$ een eindig rooster is die aan Conditie (D) voldoet. Voor elke $i \in \{1, \dots, n\}$, stel $a_i = |A_i|$ en noem b_i een natuurlijk getal die voldoet aan $1 \leq b_i \leq a_i$. Onderstel eveneens dat $f \in R[t_1, \dots, t_n]$ een niet-nul polynoom is die voldoet aan $\deg_{t_i} f \leq a_i - b_i$ voor alle $i \in \{1, \dots, n\}$. Definieer $\mathcal{U}_A = \{x \in A \mid f(x) \neq 0\}$. Als $N \in \mathbb{N}$ zodat $\sum b_i \leq N \leq \sum a_i$, dan stelt $\mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; N)$ de minimale waarde voor bereikt door alle producten $y_1 y_2 \cdots y_n$, waarbij $\sum y_i = N$ en elke y_i behoort tot het interval $[b_i, a_i]$. Dan geldt*

$$|\mathcal{U}_A| \geq \mathbf{m}(a_1, \dots, a_n; b_1, \dots, b_n; \sum_{i=1}^n a_i - \deg f).$$

Bovendien kan voor zulke $R, A_1, \dots, A_n, b_1, \dots, b_n$ een polynoom f gevonden worden waarvoor de ondergrens bereikt wordt.

De volgende gevolgen van het veralgemeend Alon-Füredi theorema betreffen nieuwe resultaten in de eindige meetkunde. Een *partiële cover* van $\text{PG}(n, q)$ wordt gedefinieerd als een collectie \mathcal{H} van hypervlakken die niet alle punten bedekken. De punten die niet bedekt worden, worden *gaten* genoemd.

Theorem A.2.3. *Onderstel dat \mathcal{H} een partiële cover van $\text{PG}(n, q)$ is die $k \geq 1$ hypervlakken bevat. Dan heeft \mathcal{H} tenminste $\mathbf{m}(q, \dots, q; nq - k + 1)$ gaten.*

Theorem A.2.4. *Als S een verzameling van k punten in $\text{AG}(n, q)$ is, dan bestaan er tenminste $\mathbf{m}(q, \dots, q; nq - k + 1) - 1$ hypervlakken van $\text{AG}(n, q)$ die disjunct zijn met S .*

De volgende versie van het Schwartz-Zippel lemma (die rekening houdt met multipliciteiten) werd eveneens aangetoond.

Theorem A.2.5. *Onderstel dat $A = \prod_{i=1}^n A_i \subseteq R^n$ een eindig rooster is met $|A_1| \geq \dots \geq |A_n|$ die aan Conditie (D) voldoet. Voor elk niet-nul polynoom $f \in R[t_1, \dots, t_n]$ geldt dan dat*

$$\sum_{x \in A} m(f, x) \leq (\deg f) \prod_{i=1}^{n-1} |A_i|,$$

waarbij $m(f, x)$ de multipliciteit van f in x voorstelt.

Bibliography

- [1] N. Alon. Combinatorial nullstellensatz. *Combin. Probab. Comput.*, 8:7–30, 1999. (Cited on pages 106 and 112.)
- [2] N. Alon, S. Friedland, and G. Kalai. Regular subgraphs of almost regular graphs. *J. Combin. Theory Ser. B*, 37(1):79–91, 1984. (Cited on page 106.)
- [3] N. Alon and Z. Füredi. Covering the cube by affine hyperplanes. *European J. Combin.*, 14(2):79–83, 1993. (Cited on pages 106, 111, 115, 129, and 130.)
- [4] N. Alon, M. B. Nathanson, and I. Ruzsa. The polynomial method and restricted sums of congruence classes. *J. Number Theory*, 56(2):404–417, 1996. (Cited on page 106.)
- [5] N. Alon and M. Tarsi. Colorings and orientations of graphs. *Combinatorica*, 12:125–134, 1992. (Cited on pages 106 and 111.)
- [6] V. Arvind, P. Mukhopadhyay, and S. Srinivasan. New results on noncommutative and commutative polynomial identity testing. *Comput. Complexity*, 19(4):521–558, 2010. (Cited on page 122.)
- [7] M. Aschbacher. Flag structures on Tits geometries. *Geom. Dedicata*, 14(21–32), 1983. (Cited on page 65.)
- [8] L. Babai and P. Frankl. *Linear Algebra Methods in Combinatorics: With Applications to Geometry and Computer Science*. Department of Computer Science, University of Chicago, 1992. (Cited on page 105.)
- [9] C. Bailey and R. B. Richter. Sum zero (mod n), size n subsets of integers. *Amer. Math. Monthly*, 96(3):240–242, 1989. (Cited on page 106.)
- [10] S. Ball. On sets of vectors of a finite vector space in which every subset of basis size is a basis. *J. Eur. Math. Soc.*, 14:733–748, 2012. (Cited on page 107.)
- [11] S. Ball. The polynomial method in Galois geometries. In *Current research topics in Galois geometry*, chapter 5, pages 105–130. Nova Science Publishers, 2012. (Cited on page 107.)
- [12] S. Ball. *Finite Geometry and Combinatorial Applications*, volume 82 of *London Mathematical Society Student Texts*. Cambridge University Press, 2015. (Cited on pages 5 and 8.)
- [13] S. Ball and O. Serra. Punctured combinatorial nullstellensätze. *Combinatorica*, 29(5):511–522, 2009. (Cited on pages 106 and 113.)
- [14] S. Ball and O. Serra. Erratum: Punctured combinatorial nullstellensätze. *Combinatorica*, 31:377–378, 2011. (Cited on page 113.)

- [15] J. Bamberg, A. Betten, Philippe C., J. De Beule, M. Lavrauw, and M. Neunhöffer. *FinInG – Finite Incidence Geometry, Version 1.0*, 2015. (Cited on page 34.)
- [16] M. K. Bardoe. The universal embedding for the involution geometry of the Suzuki sporadic simple group. *J. Algebra*, 186:447–460, 1996. (Cited on pages 66 and 90.)
- [17] M. K. Bardoe. The universal embedding for the involution geometry of Co_1 . *J. Algebra*, 217:555–572, 1999. (Cited on pages 66 and 90.)
- [18] A. Bishnoi, P. L. Clark, A. Potukuchi, and J. R. Schmitt. On zeros of a polynomial in a finite grid. *arXiv preprint <http://arxiv.org/abs/1508.06020>*, 2015. (Cited on page xi.)
- [19] A. Bishnoi and B. De Bruyn. Characterizations of the Suzuki tower near polygons. *Des. Codes Cryptogr.*, 2016. (Cited on page xi.)
- [20] A. Bishnoi and B. De Bruyn. The $L_3(4)$ near octagon. *Submitted*, 2016. (Cited on page xi.)
- [21] A. Bishnoi and B. De Bruyn. A new near octagon and the Suzuki tower. *Electron. J. Combin.*, 20:#P2.35, 2016. (Cited on page xi.)
- [22] A. Bishnoi and B. De Bruyn. On generalized hexagons of order $(3, t)$ and $(4, t)$ containing a subhexagon. *To appear in European J. Combin. arXiv preprint <http://arxiv.org/abs/1607.01004>*, 2016. (Cited on page xi.)
- [23] A. Bishnoi and B. De Bruyn. On semi-finite hexagons of order $(2, t)$ containing a subhexagon. *Ann. Comb.*, 23:433–452, 2016. (Cited on page xi.)
- [24] A. Bishnoi and F. Ihringer. Some non-existence results for distance- j ovoids in small generalized polygons. *arXiv preprint <http://arxiv.org/abs/1606.07288>*, 2016. (Cited on page xi.)
- [25] A. Blokhuis. Polynomials in finite geometries and combinatorics. In *Surveys in combinatorics*, volume 187, pages 35–52. Cambridge Univ. Press Cambridge, 1993. (Cited on page 106.)
- [26] A. Blokhuis and A. E. Brouwer. Blocking sets in desarguesian projective planes. *Bull. London Math. Soc.*, 18:132 – 134, 1986. (Cited on page 133.)
- [27] A. Blokhuis, A. E. Brouwer, and T. Szőnyi. Covering all points except one. *J. Algebraic Combin.*, 32:59–66, 2010. (Cited on page 112.)
- [28] A. Blokhuis and F. Mazzocca. The finite field Kakeya problem. In *Building Bridges*, pages 205–218. Springer, 2008. (Cited on page 107.)
- [29] D. Brink. Chevalley’s theorem with restricted variables. *Combinatorica*, 31(1):127–130, 2011. (Cited on pages 106 and 112.)
- [30] A. E. Brouwer. The Cohen-Tits near octagon on 315 points. <http://www.win.tue.nl/~aeb/graphs/HJ315.html>. Accessed 2016-04-20. (Cited on pages 12 and 92.)
- [31] A. E. Brouwer. A slowly growing collection of graph descriptions. <http://www.win.tue.nl/~aeb/graphs/index.html>. Accessed 2016-08-31. (Cited on page 86.)

-
- [32] A. E. Brouwer. A non-degenerate generalized quadrangle with lines of size four is finite. In *Advances in finite geometries and designs, Proc. 3rd Isle of Thorns Conf., Chelwood Gate/UK (1990)*, pages 47–49. Oxford University Press, 1991. (Cited on page 47.)
 - [33] A. E. Brouwer. Klabotscript, Liber Amicorum voor Henk van Tilborg, chapter Een paar gelijkmatig verdeelde codes, pages 9–14., October 2011. (Cited on page 31.)
 - [34] A. E. Brouwer, A. M. Cohen, and A. Neumaier. *Distance-regular graphs*, volume 18 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1989. (Cited on pages 11, 12, 66, and 90.)
 - [35] A. E. Brouwer, A. M. Cohen, H. A. Wilbrink, and J. I. Hall. Near polygons and Fischer spaces. *Geom. Dedicata*, 49:349–368, 1994. (Cited on page 16.)
 - [36] A. E. Brouwer, D. G. Fon-Der-Flaass, and S. V. Shpectorov. Locally co-Heawood graphs. In *Finite geometry and combinatorics (Deinze, 1992)*, volume 191 of *London Math. Soc. Lecture Note Ser.*, pages 59–68. Cambridge Univ. Press, Cambridge, 1993. (Cited on page 90.)
 - [37] A. E. Brouwer and A. Schrijver. The blocking number of an affine space. *J. Comb. Theory Ser. A*, 24(2):251–253, 1978. (Cited on pages 106 and 133.)
 - [38] A. E. Brouwer and H. A. Wilbrink. The structure of near polygons with quads. *Geom. Dedicata*, 14(2):145–176, 1983. (Cited on pages 13 and 14.)
 - [39] A. A. Bruen and J. C. Fisher. The Jamison method in Galois geometries. *Des. Codes Cryptogr.*, 1(3):199–205, 1991. (Cited on page 106.)
 - [40] F. Buekenhout and H. Van Maldeghem. Finite distance-transitive generalized polygons. *Geom. Dedicata*, 52(1):41–51, 1994. (Cited on pages 11 and 34.)
 - [41] P. J. Cameron. Orbits of permutation groups on unordered sets, II. *J. London Math. Soc.*, 2:249–264, 1981. (Cited on pages 14 and 47.)
 - [42] P. J. Cameron. Dual polar spaces. *Geom. dedicata*, 12(1):75–85, 1982. (Cited on pages 11 and 15.)
 - [43] P. J. Cameron and P. A. Kazanidis. Cores of symmetric graphs. *J. Aust. Math. Soc.*, 85(2):145–154, 2008. (Cited on page 31.)
 - [44] P. J. Cameron, J. A. Thas, and S. E. Payne. Polarities of generalized hexagons and perfect codes. *Geom. Dedicata*, 5(4):525–528, 1976. (Cited on page 31.)
 - [45] C. Carvalho. On the second Hamming weight of some Reed–Muller type codes. *Finite Fields Appl.*, 24:88–94, 2013. (Cited on page 115.)
 - [46] G. Cherlin. Locally finite generalized quadrangles with at most five points per line. *Discrete Math.*, 291(1):73–79, 2005. (Cited on page 47.)
 - [47] C. Chevalley. Démonstration d’une hypothèse de M. Artin. *Abh. Math. Sem. Univ. Hamburg*, 11:73–75, 1935. (Cited on pages 106, 111, and 112.)
 - [48] P. L. Clark. Covering numbers in linear algebra. *Amer. Math. Monthly*, 119(1):65–67, 2012. (Cited on page 132.)

- [49] P. L. Clark. The Combinatorial Nullstellensätze Revisited. *Electron. J. Combin.*, 21:#P4.15, 2014. (Cited on pages 106, 109, 111, and 112.)
- [50] P. L. Clark. Fattening up Warning’s Second Theorem. *arXiv preprint <http://arxiv.org/abs/1506.06743>*, 2015. (Cited on pages 113, 115, and 117.)
- [51] P. L. Clark, A. Forrow, and J. R. Schmitt. Warning’s Second Theorem with Restricted Variables. *Combinatorica*, 2016. (Cited on pages 106, 116, 117, and 118.)
- [52] A. M. Cohen. Geometries originating from certain distance-regular graphs. In *Finite geometries and designs*, London Math. Soc. Lecture Note Ser. 49, pages 81–87. Cambridge Univ. Press, 1981. (Cited on pages 12, 63, 65, and 135.)
- [53] A. M. Cohen and J. Tits. On generalized hexagons and a near octagon whose lines have three points. *European J. Combin.*, 6:13–27, 1985. (Cited on pages 11, 13, 16, 47, 52, 53, 54, and 80.)
- [54] J. H. Conway, R. T. Curtis, S. P. Norton, and R. A. Parker. *Atlas of finite groups*. Oxford University Press, 1985. Online version available at <http://brauer.maths.qmul.ac.uk/Atlas/v3/>. (Cited on pages 34, 63, 65, 67, 69, and 77.)
- [55] J. De Beule and L. Storme. *Current research topics on Galois geometry*. Nova Science Publishers, 2012. (Cited on page 133.)
- [56] B. De Bruyn. New near polygons from Hermitian varieties. *Bull. Belg. Math. Soc. Simon Stevin*, 10:561–577, 2003. (Cited on page 65.)
- [57] B. De Bruyn. Distance- j -ovoids in regular near $2d$ -gons. *Graphs Combin.*, 22(2):203–216, 2006. (Cited on page 32.)
- [58] B. De Bruyn. *Near polygons*. Frontiers in Mathematics. Birkhäuser, Basel, 2006. (Cited on pages 11, 12, 13, 14, 16, and 17.)
- [59] B. De Bruyn. The valuations of the near polygon \mathbb{G}_n . *Electron. J. Combin.*, 16:#R137, 2009. (Cited on pages 17, 21, and 137.)
- [60] B. De Bruyn. Polygonal valuations. *Discrete Math.*, 313:84–93, 2013. (Cited on pages 55, 56, and 57.)
- [61] B. De Bruyn. The use of valuations for classifying point-line geometries. In *Groups of exceptional type, Coxeter groups and related geometries*, Springer Proc. Math. Stat. 82, pages 27–40. Springer, 2014. (Cited on pages 17 and 57.)
- [62] B. De Bruyn. The uniqueness of a certain generalized octagon of order $(2, 4)$. *Discrete Math.*, 338(12):2125–2142, 2015. (Cited on pages 17, 38, and 57.)
- [63] B. De Bruyn. *An introduction to Incidence Geometry*. Frontiers in Mathematics. Birkhäuser, Basel, 2016. (Cited on pages 5, 6, 8, 10, 11, 12, 15, and 18.)
- [64] B. De Bruyn and P. Vandecasteele. Valuations of near polygons. *Glasg. Math. J.*, 47(02):347–361, 2005. (Cited on pages 17, 34, and 137.)
- [65] B. De Bruyn and P. Vandecasteele. The classification of the slim dense near octagons. *European J. Combin.*, 28(1):410–428, 2007. (Cited on page 17.)
- [66] B. De Bruyn and F. Vanhove. Inequalities for regular near polygons, with applications to m -ovoids. *European J. Combin.*, 34(2):522–538, 2013. (Cited on page 59.)

-
- [67] T. De Medts and H. Van Maldeghem. The uniqueness of a generalized hexagon of order 3 containing a subhexagon of order $(1, 3)$. *Discrete Math.*, 309(4):714–720, 2009. (Cited on pages 17 and 57.)
 - [68] A. De Wispelaere. *Ovoids and Spreads of Finite Classical Generalized Hexagons and Applications*. PhD thesis, Universiteit Gent, 2006. (Cited on page 31.)
 - [69] A. De Wispelaere and H. Van Maldeghem. A distance-2-spread of the generalized hexagon $H(3)$. *Ann. Comb.*, 8(2):133–154, 2004. (Cited on page 31.)
 - [70] A. De Wispelaere and H. Van Maldeghem. On the Hall-Janko graph with 100 vertices and the near-octagon of order $(2, 4)$. *Contrib. Discrete Math.*, 4:37–58, 2009. (Cited on pages 64 and 78.)
 - [71] P. Delsarte, J. M. Goethals, and F. J. Mac Williams. On generalized Reed-Muller codes and their relatives. *Information and control*, 16(5):403–442, 1970. (Cited on page 129.)
 - [72] P. Dembowski. *Finite Geometries: Reprint of the 1968 edition*. Classics in Mathematics. Springer, 1997. (Cited on page 5.)
 - [73] R. A. DeMillo and R. J. Lipton. A probabilistic remark on algebraic program testing. *Inform. Process. Lett.*, 7(4):193–195, 1978. (Cited on page 121.)
 - [74] The Sage Developers. *Sage Mathematics Software (Version 6.3)*, 2014. <http://www.sagemath.org>. (Cited on pages 31 and 33.)
 - [75] S. Dodunekov, L. Storme, and G. Van de Voorde. Partial Covers of $PG(n, q)$. *European J. Combin.*, 31:1611–1616, 2010. (Cited on pages 132 and 134.)
 - [76] Z. Dvir. On the size of Kakeya sets in finite fields. *J. Amer. Math. Soc.*, 22(4):1093–1097, 2009. (Cited on page 107.)
 - [77] Z. Dvir. Incidence theorems and their applications. *Found. Trends Theor. Comput.*, 6:257–393, 2012. (Cited on page 107.)
 - [78] Z. Dvir, S. Kopparty, S. Saraf, and M. Sudan. Extensions to the Method of Multiplicities, with Applications to Kakeya Sets and Mergers. *SIAM J. Comput.*, 42:2305–2328, 2013. (Cited on pages 116, 123, 124, 125, 126, and 127.)
 - [79] P. Erdős, A. Ginzburg, and A. Ziv. Theorem in the additive number theory. *Bull. Res. Council Israel*, 10:41–43, 1961. (Cited on page 106.)
 - [80] W. Feit and G. Higman. The nonexistence of certain generalized polygons. *J. algebra*, 1(2):114–131, 1964. (Cited on page 9.)
 - [81] D. Frohardt and P. Johnson. Geometric hyperplanes in generalized hexagons of order $(2, 2)$. *Comm. Algebra*, 22(3):773–797, 1994. (Cited on page 38.)
 - [82] The GAP Group. *GAP – Groups, Algorithms, and Programming, Version 4.8.3*, 2016. (Cited on pages 31, 34, and 69.)
 - [83] O. Geil. On the second weight of generalized Reed-Muller codes. *Des. Codes Cryptogr.*, 48(3):323–330, 2008. (Cited on page 115.)
 - [84] O. Geil and C. Thomsen. Weighted Reed-Muller codes revisited. *Des. Codes Cryptogr.*, 66(1-3):195–220, 2013. (Cited on pages 115 and 127.)

- [85] C. Godsil and G. Royle. *Algebraic graph theory*. Graduate Texts in Mathematics 207. Springer-Verlag, New York, 2001. (Cited on page 6.)
- [86] L. Guth. *Polynomial Methods in Combinatorics*, volume 64 of *University Lecture Series*. American Mathematical Society, Providence, 2016. (Cited on page 107.)
- [87] L. Guth and N. H. Katz. On the Erdős distinct distances problem in the plane. *Ann. of Math.*, 181(1):155–190, 2015. (Cited on page 107.)
- [88] W. Haemers and C. Roos. An inequality for generalized hexagons. *Geom. Dedicata*, 10(1):219–222, 1981. (Cited on page 10.)
- [89] H. Hasse. Theorie der höheren Differentiale in einem algebraischen Funktorenkörper mit vollkommenem Konstantenkörper bei beliebiger Charakteristik. *J. Reine Angew. Math.*, 175:50–54, 1936. (Cited on page 124.)
- [90] D. G. Higman. Partial geometries, generalized quadrangles and strongly regular graphs. *Atti del convegno di geometria combinatoria e sue applicazioni (Univ. Perugia, Perugia, 1970)*, pages 263–293, 1971. (Cited on page 10.)
- [91] D. G. Higman. Invariant relations, coherent configurations and generalized polygons. In *Combinatorics: Proceedings of the NATO Advanced Study Institute*, volume 16 of *NATO Advanced Study Institutes Series*, pages 347–363. Springer, 1975. (Cited on page 10.)
- [92] J. W. P. Hirschfeld and J. A. Thas. *General galois geometries*. Springer, 1991. (Cited on page 5.)
- [93] R. E. Jamison. Covering finite fields with cosets of subspaces. *J. Comb. Theory Ser. A*, 22(3):253–266, 1977. (Cited on pages 105 and 133.)
- [94] W. M. Kantor. Generalized polygons, scabs and gabs. In *Buildings and the Geometry of Diagrams*, pages 79–158. Springer, 1986. (Cited on page 5.)
- [95] R. N. Karasev and F. V. Petrov. Partitions of nonzero elements of a finite field into pairs. *Israel J. Math.*, 192(1):143–156, 2012. (Cited on page 106.)
- [96] G. Károlyi and Z. Nagy. A simple proof of the Zeilberger–Bressoud q -Dyson theorem. *Proc. Amer. Math. Soc.*, 142(9):3007–3011, 2014. (Cited on page 106.)
- [97] T. Kasami, S. Lin, and W. W. Peterson. Generalized Reed-Muller codes. *Electron. Commun. Japan*, 51(3):96–104, 1968. (Cited on page 129.)
- [98] D. E. Knuth. Dancing links. In *Millennial Perspectives in Computer Science, Cornerstones of Computing*, pages 187–214. Palgrave Macmillan, 2000. (Cited on page 33.)
- [99] P. Komjáth. Partitions of vector spaces. *Periodica Mathematica Hungarica*, 28:187–193, 1994. (Cited on page 115.)
- [100] T. H. Koornwinder. A note on the absolute bound for systems of lines. In *Indag. Math. (Proceedings)*, volume 79, pages 152–153. Elsevier, 1976. (Cited on page 105.)
- [101] D. Leemans. A family of geometries related to the Suzuki tower. *Comm. Algebra*, 33:2201–2217, 2005. (Cited on pages 89 and 90.)

-
- [102] P. W. H. Lemmens and J. J. Seidel. Equiangular lines. *J. Algebra*, 24(3):494–512, 1973. (Cited on page 105.)
 - [103] R. Lidl and H. Niederreiter. *Finite fields*, volume 20. Cambridge university press, 1997. (Cited on page 106.)
 - [104] J. H. Van Lint. *Introduction to coding theory*. Graduate Texts in Mathematics 86. Springer-Verlag, Berlin, third edition, 1999. (Cited on page 128.)
 - [105] S. Linton. Finding the smallest image of a set. In *Proceedings of the 2004 International Symposium on Symbolic and Algebraic Computation*, ISSAC '04, pages 229–234, New York, NY, USA, 2004. ACM. (Cited on page 37.)
 - [106] R. J. Lipton. The curious history of the Schwartz-Zippel lemma. <https://rjlipton.wordpress.com/2009/11/30>. Accessed 2016-04-22. (Cited on page 121.)
 - [107] H. H. López, C. Rentería-Márquez, and R. H. Villarreal. Affine cartesian codes. *Des. Codes and Cryptogr.*, 71(1):5–19, 2014. (Cited on pages 115, 129, and 130.)
 - [108] J. Matoušek. *Thirty-three miniatures*. Student Mathematical Library. American Mathematical Society, 2010. (Cited on page 105.)
 - [109] K. Metsch. How many s -subspaces must miss a point set in $PG(d, q)$. *J. Geom.*, 86:154–164, 2006. (Cited on page 133.)
 - [110] E. Moorhouse. Generalised Polygons of Small Order. <http://www.uwo.edu/moorhouse/pub/genpoly/>. Accessed 2016-04-20. (Cited on page 36.)
 - [111] D. E. Muller. Application of Boolean algebra to switching circuit design and to error detection. *IRE Tran. on Electronic Computers EC-3*, 3:6–12, 1954. (Cited on page 129.)
 - [112] A. Neumaier. Rectagraphs, diagrams, and Suzuki’s sporadic simple group. In *Algebraic and geometric combinatorics*, volume 65 of *North-Holland Math. Stud.*, pages 305–318. North-Holland, Amsterdam, 1982. (Cited on page 90.)
 - [113] A. Offer and H. Van Maldeghem. Distance- j ovoids and related structures in generalized polygons. *Discrete Math.*, 294(1-2):147–160, 2005. (Cited on pages 31 and 32.)
 - [114] O. Ore. über höhere Kongruenzen. *Norsk Mat. Forenings Skrifter Ser. I*, 7:15 pages, 1922. (Cited on page 123.)
 - [115] D. Pasechnik. Geometric characterization of graphs from the Suzuki chain. *European J. Combin.*, 14:491–499, 1993. (Cited on pages 90 and 91.)
 - [116] S. E. Payne and J. A. Thas. *Finite generalized quadrangles*. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, second edition, 2009. (Cited on pages 8, 10, 11, and 14.)
 - [117] C. Pech and S. Reichard. Enumerating set orbits. In *Algorithmic algebraic combinatorics and Gröbner bases*, pages 137–150. Springer, Berlin, 2009. (Cited on pages 31 and 43.)
 - [118] I. S. Reed. A class of multiple-error-correcting codes and the decoding scheme. *IRE Trans. Information Theory PGIT*, 4:38–49, 1954. (Cited on page 129.)

- [119] M. A. Ronan. A geometric characterization of Moufang hexagons. *Invent. Math.*, 57(3):227–262, 1980. (Cited on page 11.)
- [120] M. A. Ronan. A combinatorial characterization of the dual Moufang hexagons. *Geom. Dedicata*, 11(1):61–67, 1981. (Cited on pages 34, 77, and 91.)
- [121] U. Schauz. Algebraically solvable problems: describing polynomials as equivalent to explicit solutions. *Electron. J. Combin.*, 15(1):P10–35, 2008. (Cited on pages 109, 111, and 112.)
- [122] J. T. Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *J. Assoc. Comput. Mach.*, 4:701–717, 1980. (Cited on page 121.)
- [123] B. Segre. Ovals in a finite projective plane. *Canad. J. Math.*, 7:414–416, 1955. (Cited on page 107.)
- [124] A. Shpilka and A. Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3–4):207–388, 2010. (Cited on page 121.)
- [125] E. E. Shult. *Points and Lines: Characterizing the Classical Geometries*. Springer, 2011. (Cited on page 5.)
- [126] E. E. Shult and A. Yanushka. Near n -gons and line systems. *Geom. Dedicata*, 9:1–72, 1980. (Cited on pages 11, 13, 14, 19, and 65.)
- [127] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. Published electronically at <https://oeis.org>. (Cited on page 44.)
- [128] L. H. Soicher. On simplicial complexes related to the Suzuki sequence graphs. In *Groups, combinatorics and geometry (Durham, 1990)*, volume 165 of *London Math. Soc. Lecture Note*, pages 240–248. Cambridge, 1992. (Cited on page 90.)
- [129] L. H. Soicher. Three new distance-regular graphs. *European J. Combin.*, 14:501–505, 1993. (Cited on pages 65, 85, and 90.)
- [130] L. H. Soicher. GRAPE, GRaph algorithms using PERmutation groups, Version 4.6.1. <http://www.maths.qmul.ac.uk/~leonard/grape/>, May 2012. Refereed GAP package. (Cited on page 37.)
- [131] L. H. Soicher. The uniqueness of a distance-regular graph with intersection array $\{32, 27, 8, 1; 1, 4, 27, 32\}$ and related results. *Des. Codes Cryptogr.*, 2016. (Cited on page 85.)
- [132] M. Suzuki. A simple group of order 448, 345, 497, 600. In *Theory of Finite Groups (Symposium, Harvard Univ., Cambridge, Mass., 1968)*, pages 113–119. Benjamin, New York, 1969. (Cited on pages 89 and 91.)
- [133] T. Tao. Algebraic combinatorial geometry: the polynomial method in arithmetic combinatorics, incidence combinatorics, and number theory. *EMS Surveys in Mathematical Sciences*, 1:1–46, 2014. (Cited on page 107.)
- [134] T. Tao and V. H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2006. (Cited on page 106.)

-
- [135] J. A. Thas. Ovoids and spreads of finite classical polar spaces. *Geom. Dedicata*, 10:135–143, 1981. (Cited on page 31.)
 - [136] J. Tits. Sur la trialité et certains groupes qui s'en déduisent. *Inst. Hautes Études Sci. Publ. Math.*, 2:13–60, 1959. (Cited on pages 5 and 9.)
 - [137] J. Tits. Moufang octagons and the Ree groups of type 2F_4 . *Amer. J. Math.*, 105:539–594, 1983. (Cited on page 9.)
 - [138] H. Van Maldeghem. *Generalized Polygons*. Birkhäuser Basel, 1998. (Cited on pages 5, 7, 9, 11, 45, 47, and 136.)
 - [139] E. Warning. Bemerkung zur vorstehenden Arbeit von Herrn Chevalley. *Abh. Math. Sem. Hamburg*, 11:76–83, 1935. (Cited on pages 112 and 116.)
 - [140] R. A. Wilson. *The finite simple groups*. Graduate Texts in Mathematics 251. Springer-Verlag, London, 2009. (Cited on page 89.)
 - [141] S. Yoshiara. A lattice theoretical construction of a GAB of the Suzuki sporadic simple group. *J. Algebra*, 112:198–239, 1988. (Cited on page 90.)
 - [142] S. Yoshiara. On the geometry of the Hall-Janko group on 315 points. *Geom. Dedicata*, 32:173–201, 1989. (Cited on pages 63 and 66.)
 - [143] R. Zippel. Probabilistic algorithms for sparse polynomials. In *Lect. Notes in Comput. Sci. 72, Proc. EUROSAM 79*. Springer, 1979. (Cited on page 121.)

